

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2002, 08:59:13 : Search time 14 Seconds
(without alignments)
108.170 Million cell updates/sec

Title: US-09-640-636A-4

Perfect score: 338
Sequence: 1 RGDHKLHNPALQDPLQLSL.....QSDGVLGHLQRPCHVRLGTA 62

Scoring table:

BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued_patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PTUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfilest.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	64	18.9	145	2	US-08-832-535-11
2	64	18.9	146	2	US-08-791-522-3
3	64	18.9	146	3	US-08-744-138-3
4	64	18.9	146	3	US-09-019-485-4
5	64	18.9	146	3	US-09-314-777-3
6	64	18.9	146	4	US-09-431-480-6
7	64	18.9	146	4	US-09-617-302-6
8	64	18.9	146	4	US-09-241-376-3
9	64	18.9	146	5	PCT-US95-07135-9
10	64	18.9	146	6	5432264-6
11	63.5	18.8	1213	4	US-08-413-814-79
12	62	18.3	617	1	US-08-137-614A-26
13	62	18.3	637	1	US-08-072-064-1
14	62	18.3	637	3	US-08-072-064-4
15	62	18.3	637	3	US-08-072-064-6
16	62	18.3	637	3	US-08-072-064-8
17	62	18.3	637	5	PCT-US92-08558-1
18	61.5	18.2	469	3	US-08-753-007A-8
19	61.5	18.2	469	3	US-09-398-496-8
20	59	17.5	900	2	US-08-813-940-4
21	58	17.2	39	6	5432264-8
22	58	17.2	554	4	US-08-462-467B-22
23	58	17.2	1399	4	US-08-462-467B-14
24	57	16.9	425	4	US-08-462-467B-16
25	57	16.9	649	4	US-08-462-467B-20
26	57	16.9	1495	4	US-08-462-467B-12
27	56.5	16.7	910	4	US-08-460-269C-2

28	56	16.6	349	4	US-08-462-467B-18	Sequence 18, Appl
29	55.5	16.4	855	4	US-09-813-819-2	Sequence 2, Appl
30	55.5	16.4	855	4	US-09-920-048-2	Sequence 2, Appl
31	55	16.3	183	6	5168049-4	Patent No. 5168049
32	55	16.3	412	4	US-09-027-064-4	Sequence 4, Appl
33	55	16.3	412	4	US-09-271-815-4	Sequence 4, Appl
34	55	16.3	677	4	US-08-464-601-2	Sequence 2, Appl
35	55	16.3	677	4	US-09-210-993A-2	Sequence 2, Appl
36	54.5	16.1	384	1	US-08-221-750A-11	Sequence 11, Appl
37	54.5	16.1	472	2	US-08-459-346-17	Sequence 17, Appl
38	54.5	16.1	472	2	US-08-889-419-17	Sequence 17, Appl
39	54.5	16.1	472	5	PCT-US93-07189-17	Sequence 17, Appl
40	54.5	16.1	898	2	US-08-474-379C-12	Sequence 12, Appl
41	54.5	16.1	898	3	US-09-146-249A-12	Sequence 12, Appl
42	54.5	16.1	898	3	US-08-205-188B-12	Sequence 12, Appl
43	54.5	16.1	900	1	US-07-688-352C-12	Sequence 12, Appl
44	54	16.0	262	4	US-09-199-657A-293	Sequence 293, App
45	54	16.0	472	4	US-08-976-255-17	Sequence 17, Appl

ALIGNMENTS

RESULT 1
US-08-832-535-11
; Sequence 11, Application US/08832535
; Patent No. 5919658
; GENERAL INFORMATION:
; APPLICANT: NI, JIAN
; APPLICANT: LI, HAODONG
; APPLICANT: YU, GUO-LIANG
; APPLICANT: GENTZ, REINER L
; TITLE OF INVENTION: HUMAN CYSTATIN F
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: HUMAN GENOME SCIENCES, INC.
; STREET: 9410 KEY WEST AVENUE
; CITY: ROCKVILLE
; STATE: MD
; COUNTRY: US
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/832,535
; FILING DATE: 03-APR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: KIMBALL, PAUL C.
; REGISTRATION NUMBER: 34,610
; REFERENCE/DOCKET NUMBER: PE265
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (201) 994-1700
; TELEFAX: (201) 994-1744
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 145 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-832-535-11

Query Match 18.9%: Score 64; DB 2; Length 145;
Best Local Similarity 42.9%: Pred. No. 0.94;
Matches 12; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

QY 24 GHAPRPPRDLGGSDGVSGVSLQHL 51
Db 26 GSPGKPRPLVGGPMADAVVEEGVRL 53

RESULT 2
US-08-791-522-3
Sequence 3, Application US/08791522
Patent No. 5935817
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Goll, Surya K.
TITLE OF INVENTION: NOVEL HUMAN CYSTATIN-LIKE
TITLE OF INVENTION: PROTEIN
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/791,522
FILING DATE: Filed Herewith
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0193 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 146 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 181387
US-08-791-522-3

Query Match 18.9%; Score 64; DB 2; Length 146;
Best Local Similarity 42.9%; Pred. No. 0.95;
Matches 12; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

OY 24 GHAPPPRDLGGGDSVQSDGVQLHL 51
1 : 1 111 : 11 : 1 : 11 : 1
Db 26 GSSPGKPPRLVGGPMDASVEEGVRRAL 53

RESULT 3
US-08-744-138-3
Sequence 3, Application US/08744138
Patent No. 6011012
GENERAL INFORMATION:
APPLICANT: Gentz, Reiner L.
APPLICANT: Ni, Jian
APPLICANT: Rosen, Craig A.
APPLICANT: Yu, Guo-Liang
TITLE OF INVENTION: Human Cystatin E
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville

STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/744,138
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PP202P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301 309 8504
TELEFAX: 301 309 8512
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 146 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Cystatin C
US-08-744-138-3

Query Match 18.9%; Score 64; DB 3; Length 146;
Best Local Similarity 42.9%; Pred. No. 0.95;
Matches 12; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

OY 24 GHAPPPRDLGGGDSVQSDGVQLHL 51
1 : 1 111 : 11 : 1 : 11 : 1
Db 26 GSSPGKPPRLVGGPMDASVEEGVRRAL 53

RESULT 4
US-09-019-485-4
Sequence 4, Application US/09019485
Patent No. 6066617
GENERAL INFORMATION:
APPLICANT: Li, Haodong
APPLICANT: Yu, Guo-Liang
APPLICANT: Gentz, Reiner
APPLICANT: Ni, Jian
TITLE OF INVENTION: Cystatin F
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: MD
COUNTRY: US
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/019,485
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Benson, Robert H.
REGISTRATION NUMBER: 30,446
REFERENCE/DOCKET NUMBER: PP265P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 3013098504

TELEFAX: 3013098439
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 146 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-019-485-4

Query Match 18.9%; Score 64; DB 3; Length 146;
Best Local Similarity 42.9%; Pred. No. 0.95;
Matches 12; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

OY 24 GHAPPPRDLGGIDSGVSDGYLQHL 51
Db 26 GSSPGKPPRLVGGPMDASVEEGBVRRAL 53

RESULT 5

US-09-314-777-3
; Sequence 3, Application US/09314777
; Patent No. 6110686
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Goll, Surya K.
; TITLE OF INVENTION: NOVEL HUMAN CYSTATIN-LIKE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/314,777
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/791,522
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0193 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 146 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 181387
US-09-314-777-3

Query Match 18.9%; Score 64; DB 3; Length 146;
Best Local Similarity 42.9%; Pred. No. 0.95;
Matches 12; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

OY 24 GHAPPPRDLGGIDSGVSDGYLQHL 51
Db 26 GSSPGKPPRLVGGPMDASVEEGBVRRAL 53

RESULT 6
US-09-431-480-6
; Sequence 6, Application US/09431480
; Patent No. 6235708
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; APPLICANT: Feldhaus, Andrew
; TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
; FILE REFERENCE: 98-72
; CURRENT APPLICATION NUMBER: US/09/431,480
; CURRENT FILING DATE: 1999-11-01
; EARLIER APPLICATION NUMBER: 60/109,217
; EARLIER FILING DATE: 1998-11-20
; EARLIER APPLICATION NUMBER: 60/156,382
; EARLIER FILING DATE: 1999-09-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 146
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-431-480-6

Query Match 18.9%; Score 64; DB 4; Length 146;
Best Local Similarity 42.9%; Pred. No. 0.95;
Matches 12; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

OY 24 GHAPPPRDLGGIDSGVSDGYLQHL 51
Db 26 GSSPGKPPRLVGGPMDASVEEGBVRRAL 53

RESULT 7

US-09-617-302-6
; Sequence 6, Application US/09617302
; Patent No. 6245529
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; APPLICANT: Feldhaus, Andrew
; TITLE OF INVENTION: TESTIS SPECIFIC CYSTATIN-LIKE PROTEIN CYSTATIN T
; FILE REFERENCE: 98-72 CI
; CURRENT APPLICATION NUMBER: US/09/617,302
; CURRENT FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 09/431,480
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: 60/109,217
; PRIOR FILING DATE: 1998-11-20
; PRIOR APPLICATION NUMBER: 60/156,382
; PRIOR FILING DATE: 1999-09-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 146
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-617-302-6

Query Match 18.9%; Score 64; DB 4; Length 146;
Best Local Similarity 42.9%; Pred. No. 0.95;
Matches 12; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

OY 24 GHAPPPRDLGGIDSGVSDGYLQHL 51
Db 26 GSSPGKPPRLVGGPMDASVEEGBVRRAL 53

RESULT 8
US-09-241-376-3
; Sequence 3, Application US/09241376
; Patent No. 6300477
; GENERAL INFORMATION:

APPLICANT: Gentz, Reiner L.
APPLICANT: Ni, Jian
APPLICANT: Rosen, Craig A.
APPLICANT: Yu, Guo-Liang
TITLE OF INVENTION: Human Cystatin E
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/241,376
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/744,138
FILING DATE: 05-NOV-1996
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/461,030
FILING DATE: 05-JUN-1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PF202P1D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301 309 8504
TELEFAX: 301 309 8512
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 146 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Cystatin C
US-09-241-376-3

Query Match 18.9%; Score 64; DB 4; Length 146;
Best Local Similarity 42.9%; Pred. No. 0.95;
Matches 12: Conservative 7; Mismatches 9; Indels 0; Gaps 0;

OY 24 GHAPPPRDLGGGDSVQSDGVQL 51
DB 26 GSSPGKPRPLVGGPMDASVEEGVRRAL 53

RESULT 9
PCT-US95-07135-9
Sequence 9, Application PC/TUS9507135
GENERAL INFORMATION:
APPLICANT: Ni, Jian
APPLICANT: Yu, Guo-Liang
APPLICANT: Gentz, Reiner
APPLICANT: Rosen, Craig
TITLE OF INVENTION: HUMAN CYSTATIN E
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESSES:
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
ADDRESS: CROCHT, STUART & OLSTEIN
STREET: 6 Becker Farm Road
CITY: Roseland

STATE: New Jersey
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/07135
FILING DATE: 05-JUN-1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: MULLINS, J.G.
REGISTRATION NUMBER: 33,073
REFERENCE/DOCKET NUMBER: 325800-400
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 146 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-07135-9

Query Match 18.9%; Score 64; DB 5; Length 146;
Best Local Similarity 42.9%; Pred. No. 0.95;
Matches 12: Conservative 7; Mismatches 9; Indels 0; Gaps 0;

OY 24 GHAPPPRDLGGGDSVQSDGVQL 51
DB 26 GSSPGKPRPLVGGPMDASVEEGVRRAL 53

RESULT 10
5432264-6
Patent No. 5432264
APPLICANT: GRUBB, ANDERS; LUNDMALL, AKE; ABRAHAMSON, MAGNUS;
DALBOGE, HENRIK
TITLE OF INVENTION: RECOMBINANT 3-DES-OH-CYSTATIN C PRODUCED
BY EXPRESSION IN A PROCARYOTIC HOST CELL
NUMBER OF SEQUENCES: 8
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/929,290
FILING DATE: 13-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 440,221
FILING DATE: 21-NOV-1989
APPLICATION NUMBER: 297,198
FILING DATE: 20-MAY-1988
SEQ ID NO: 6:
LENGTH: 146
5432264-6

Query Match 18.9%; Score 64; DB 6; Length 146;
Best Local Similarity 42.9%; Pred. No. 0.95;
Matches 12: Conservative 7; Mismatches 9; Indels 0; Gaps 0;

OY 24 GHAPPPRDLGGGDSVQSDGVQL 51
DB 26 GSSPGKPRPLVGGPMDASVEEGVRRAL 53

RESULT 11
US-09-413-814-79
Sequence 79, Application US/09413814
Patent No. 6225064
GENERAL INFORMATION:
APPLICANT: Gesellschaft fuer Biotechnologische Forschung mbH
APPLICANT: Bristol-Myers Squibb, Co.

APPLICANT: Beyer, Stefan
APPLICANT: Bloecker, Helmut
APPLICANT: Brandt, Petra
APPLICANT: Cino, Paul M
APPLICANT: Dougherty, Brian A
APPLICANT: Goldberg, Steven L
APPLICANT: Hoffer, Gerhard
APPLICANT: Mueller, Joachim
APPLICANT: Reichenbach, Hans
TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polyketide or
TITLE OF INVENTION: heteropolyketide compounds
FILE REFERENCE: PCY/US 99/23535
CURRENT APPLICATION NUMBER: US/09/413,814
CURRENT FILING DATE: 1999-10-07
EARLIER APPLICATION NUMBER: DE 198 46 493.2
EARLIER FILING DATE: 1998-10-09
NUMBER OF SEQ ID NOS: 107
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 79
LENGTH: 1213
TYPE: PRT
ORGANISM: Sorangium cellulosum
US-09-413-814-79

Query Match 18.8%; Score 63.5; DB 4; Length 1213;
Best Local Similarity 35.5%; Pred. No. 13;
Matches 22; Conservative 8; Mismatches 23; Indels 9; Gaps 4;

QY 1 RGDHKL-AHRPALODPLL--OSLRPGHAPHPPRDLGGIDSGVSDGLQHLQRPCH 56
DB 375 RRDHRAARHRSGEDPLRPHARRDRGR-PRAPLRGRGGRGDRSG---REAPGR 429
QY 57 VR 58
DB 430 VR 431

RESULT 12
US-08-137-614A-26
Sequence 26, Application US/08137614A
Patent No. 5487976
GENERAL INFORMATION:
APPLICANT: Soderlund, David M.
APPLICANT: Knipple, Douglas C.
APPLICANT: Henderson, Joseph E.
TITLE OF INVENTION: Gene Encoding An Insect
TITLE OF INVENTION: Gamma-Aminobutyric Acid (GABA) Receptor Subunit
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon, Hargrave, Devans & Doyle
STREET: Clinton Square, P.O. Box 1051
CITY: Rochester
STATE: New York
COUNTRY: USA
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/137,614A
FILING DATE: 15-OCT-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Timian, Susan J.
REGISTRATION NUMBER: 34,103
REFERENCE/DOCKET NUMBER: 19603/120
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716)263-1636
TELEFAX: (716)263-1600
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:

LENGTH: 617 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-137-614A-26

Query Match 18.3%; Score 62; DB 1; Length 617;
Best Local Similarity 36.7%; Pred. No. 9.1;
Matches 18; Conservative 2; Mismatches 17; Indels 12; Gaps 2;

QY 22 PRGHAPHP-----PPDLGGIDSGVSDG-----VLQHLQRPCHVK 58
DB 514 PSHGPGPGGGGPGGGGGGGGGGPPGGGDPPEAAVPAHLHLRQKVK 562

RESULT 13
US-08-072-064-1
Sequence 1, Application US/08072064
Patent No. 6008046
GENERAL INFORMATION:
APPLICANT: FRENCH-CONSTANT, RICHARD H.
APPLICANT: JACKSON, MEYER B.
TITLE OF INVENTION: DRUG AND PESTICIDE SCREENING
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: PETER G. CARROLL
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/072,064
FILING DATE: 19930602
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 770,881
FILING DATE: 04-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: CARROLL, PETER G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: OPHD-00574
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/705-8410
TELEFAX: 415/397-8338
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 637 amino acids
TYPE: AMINO ACID
TOPOLOGY: unknown
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Drosophila melanogaster
POSITION IN GENOME:
CHROMOSOME/SEGMENT: III; polytene subregion 66F
MAP POSITION: approximately map unit 26
US-08-072-064-1

Query Match 18.3%; Score 62; DB 3; Length 637;
Best Local Similarity 36.7%; Pred. No. 9.5;
Matches 18; Conservative 2; Mismatches 17; Indels 12; Gaps 2;

QY 22 PRGHAPHP-----PPDLGGIDSGVSDG-----VLQHLQRPCHVK 58
DB 535 PSHGPGPGGGGPGGGGGGGGGGPPGGGDPPEAAVPAHLHLRQKVK 583

RESULT 14

```

US-08-072-064-4
Sequence 4, Application US/08072064
Patent No. 6008046
GENERAL INFORMATION:
APPLICANT: FFERENCH-CONSTANT, RICHARD H.
APPLICANT: JACKSON, MEYER B.
TITLE OF INVENTION: DRUG AND PESTICIDE SCREENING
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESSES:
ADDRESSEE: PETER G. CARROLL
STREET: 220 Montgomery street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version.#1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/072,064
FILING DATE: 19930602
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 770,881
FILING DATE: 04-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: CARROLL, PETER G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: OCHD-00574
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/705-8410
TELEFAX: 415/397-8338
INFORMATION FOR SEQ ID NO.: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 637 amino acids
TYPE: AMINO ACID
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-072-064-4

Query Match      18.3%; Score 62; DB 3; Length 637;
Best Local Similarity 36.7%; Pred. No. 9.5;
Matches 18; Conservative 2; Mismatches 17; Indels 12; Gaps 2;

OY    22 PRGHAPH-----PPRDLGGGLDSGVQSG-----VLQHIGRPGRH 58
DB    535 POSHDPGPGGGGPRGGGGGGGGGPPRGPEPAVPAHLIHNGKVK 563

```

```

; APPLIATION NUMBER: US/08/072,064
; FILING DATE: 19930602
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 770,881
; FILING DATE: 04-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: CARROLL, PETER G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: OPHD-00574
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/705-8410
; TELEFAX: 415/397-8338
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 637 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; US-08-072-064-6

Query Match      18.3%; Score 62; DB 3; Length 637;
Best Local Similarity 36.7%; Pred. No. 9.5;
Matches 18; Conservative 2; Mismatches 17; Indels 12; Gaps 2.

QY      22 PRGHAPHP-----PPRDLGGGLGDSGVQNSG-----VLQHLDQRPGHWK 58
       1:| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      535 FSHSPRGPGGGGPRPGGGGGGGGGGPPRGGGDPPEAAVPAHLLIHFGKK 583

```

Search completed: November 8, 2002, 09:01:19
Job time : 15 secs

RESULT 15
US-08-072-064-6
Sequence 6, Application US/08072064
Patent No. 6008046
GENERAL INFORMATION:
APPLICANT: FRENCH-CONSTANT, RICHARD H.
APPLICANT: JACKSON, MEYER B.
TITLE OF INVENTION: DRUG AND PESTICIDE SCREENING
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: PETER G. CARROLL
STREET: 220 Montgomery street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2002, 08:58:18 : Search time 19 Seconds

(without alignments)
313.555 Million cell updates/sec

Title: US-09-640-636A-4

Perfect score: 338

Sequence: 1 RGDHKLHRRPALQDPILLQSL.....QSDGVQLHQRPHVKLGTA 62

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : PIR.71.*

1: PIR1.*

2: PIR2.*

3: PIR3.*

4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	74.5	22.0	700	2 D70951	probable uvrd - My
2	64.5	19.1	642	2 T35003	probable respirato
3	64	18.9	146	1 UDHU	cytatin C precurs
4	63.5	18.8	156	2 B72592	hypothetical prote
5	63	18.6	592	2 T32402	hypothetical prote
6	62.5	18.5	253	2 PNO447	hypothetical prote
7	62.5	18.5	363	2 A72702	hypothetical prote
8	62	18.3	580	2 S72211	N-acetyl-Delta-D-gl
9	62	18.3	606	2 A41145	gamma-aminobutyric
10	61	18.0	266	2 B72532	hypothetical prote
11	60.5	17.9	465	2 A96553	probable myosinas
12	60.5	17.9	1870	2 S36771	MHC class III hist
13	60.5	17.9	1872	2 S36152	MHC class III hist
14	60.5	17.9	2142	2 B35098	MHC class III hist
15	59.5	17.6	211	2 AC3290	hypothetical prote
16	59.5	17.6	411	1 J64003	cytochrome P450 -
17	59	17.5	579	2 T24880	hypothetical prote
18	59	17.5	934	2 G70563	DNA topoisomerase
19	59	17.5	1355	2 S40022	spalt protein - fr
20	58.5	17.3	245	1 S10092	homeotic protein H
21	58.5	17.3	273	2 B69883	conserved hypotet
22	58	17.2	103	2 C72683	hypothetical prote
23	58	17.2	243	2 AH3263	exsd protein limpo
24	58	17.2	272	2 T35231	hypothetical prote
25	58	17.2	426	2 C81129	hypothetical prote
26	58	17.2	426	2 E81835	probable periplasm
27	58	17.2	503	2 G75276	carboxypeptidase -
28	58	17.2	517	2 T37225	hypothetical prote
29	58	17.2	525	1 KGHUGH	histidine-rich gly

30	58	17.2	533	2 S33744	cyclodiene insecti
31	58	17.2	646	2 A44985	dnak-type molecule
32	58	17.2	958	2 T51148	hypothetical prote
33	58	17.2	1106	2 J00405	hypothetical 119.5
34	57.5	17.0	259	2 AB3572	succinoglycan bios
35	57.5	17.0	545	1 S44486	indole-3-pyruvate
36	57	16.9	141	2 C72696	hypothetical prote
37	57	16.9	209	2 C89005	protein T24A6.3 li
38	57	16.9	271	2 AB2095	hypothetical prote
39	57	16.9	271	2 F83081	hypothetical prote
40	57	16.9	411	2 G75475	conserved hypotet
41	57	16.9	419	2 D69202	5-enolpyruvylshik
42	57	16.9	518	2 AH2336	hypothetical prote
43	57	16.9	528	2 S32593	[RNA-polymerase]-s
44	57	16.9	600	2 F70980	probable accA3 pro
45	57	16.9	680	2 T45736	hypothetical prote

ALIGNMENTS

RESULT 1

D70951 Probable uvrd - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000

C:Accession: D70951

R: Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordo

; Connor, R.; Davies, R.; Devlin, K.; Feldwell, T.; Gentles, S.; Hamlin, N.; Holroyd

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A: Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A: Title: Deciphering the biology of Mycobacterium tuberculosis from the complete gen

A: Reference number: A70500; MUID: 98295987

A: Accession: D70951

A: Status: preliminary; nucleic acid sequence not shown; translation not shown

A: Molecule type: DNA

A: Residues: 1-700 <COI>

A: Cross-references: GB:AL021646; GB:AL123456; NID:93242278; PIDN:CAA1663.1; PID:928

A: Experimental source: strain H37RV

C: Genetics:

A: Gene: uvrd2

C: Superfamily: helicase II

Query Match

Best Local Similarity

Matches 18; Conservative

Mismatches 9; Indels 1; Gaps 1;

QY 3 DHKLHRRPALQDP-LLQSLRPHGHAPHP 30

DB 463 DDELAORPALQDPGLLAELRRRAEHRHP 491

RESULT 2

T35003

Probable respiratory chain oxidoreductase - Streptomyces coelicolor

C: Species: Streptomyces coelicolor

C: Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 21-Jan-2000

C: Accession: T35003

R: Seeger, K.J.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream,

submitted to the EMBL Data Library, December 1998

A: Reference number: Z21564

A: Accession: T35003

A: Status: preliminary; translated from GB/EMBL/DBJ

A: Molecule type: DNA

A: Residues: 1-642 <SEP>

A: Cross-references: EMBL:AL034443; PIDN:CAA22365.1; GSDB:CN00070; SCOEDB:SC4B5.11c

A: Experimental source: strain A5(2)

C: Genetics:

A: Gene: SCOEDB:SC4B5.11c

C: Superfamily: formate dehydrogenase

Query Match

19.1%; Score 64.5; DB 2; Length 642;

RESULT 4

B72592
hypothetical protein APE1206 - Aeropyrum pernix (strain K1)

C:Species: Aeropyrum pernix

C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999

C:Accession: B72592

R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K

DNA Res. 6, 83-101, 1999

A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum pernix

A:Reference number: A72450; MUID:99310339

A:Accession: B72592

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-156 <KAW>

A:Cross-references: DDBJ:AP000061; NID:g5104821; PIDN:BAAB0192.1; PID:d1043978; PID:g510

A:Experimental source: strain K1

C:Genetics:

A:Gene: APE1206

Query Match

Best Local Similarity

Matches 17; Conservative

18.8%; Score 63.5; DB 2; Length 156;

37.0%; Pred. No. 5;

Mismatches 21; Indels 1; Gaps 1;

OY

14

DELQSLRPRGHAPRRPRDLGGSGVQSD-GVLOHLQRRGHV 58

DB

67

NELQSPRRGLEPPYTPRVGHTLGGYPAELPVLEHNVPGDIQ 112

RESULT 5

T32402

hypothetical protein R148.5 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

C:Accession: T32402

R:Lee, T.T.; Kemp, K.; Schaefer, P.

submitted to the EMBL data library, September 1997

A:Description: The sequence of C. elegans cosmid R148.

A:Reference number: Z2161

A:Accession: T32402

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-592 <LEW>

A:Cross-references: EMBL:AF025467; PIDN:AA071039.1; GSPDB:GN00021; CESP:R148.5

A:Experimental source: strain Bristol N2; clone R148

C:Genetics:

A:Gene: CESP:R148.5

A:Map position: 3

A:Introns: 114/2; 236/1; 307/1; 533/2; 561/3

Query Match

Best Local Similarity

Matches 14; Conservative

18.6%; Score 63; DB 2; Length 592;

37.8%; Pred. No. 25;

Mismatches 17; Indels 2; Gaps 1;

OY

4

HKLARPAALQDPLQSL--RRGHAPRRPRDLGGGL 38

DB

344

HPMHHPVAMPPPPGFPFPPRRGHPPPPHFMGRGM 380

RESULT 6

PNO447

hypothetical protein 253 - Streptomyces griseus plasmid pSG1 (fragment)

C:Species: Streptomyces griseus

C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 07-May-1999

C:Accession: PNO447

R:Bar-Nir, D.; Cohen, A.; Goedeke, M.E.

Gene 122, 71-76, 1992

A:Title: tRNA-ser sequence are involved in the excision of Streptomyces griseus plasmid

A:Reference number: PNO447; MUID:93083997

A:Accession: PNO447

A:Molecule type: DNA

A:Residues: 1-253 <BAR>
A:Cross-references: GB:M86370
C:Genetics:
A:Genome: plasmid

Query Match

Best Local Similarity

Matches 24; Conservative

18.5%; Score 62.5; DB 2; Length 253;

32.9%; Pred. No. 11;

Mismatches 21; Indels 21; Gaps 5;

OY

1

RGDHLAH---RPAALQDPLQSLRPRG--HAPRRP-----PPDLGGGLDSGVQ 44

DB

184

RGNRHNPCSPRPALQGPL---VKGASVHPPTGRLSPSRWSPOSGLASRGVSQ 239

OY

45

DGVLOLQRRGHV 57

DB

240

FSLSEALIS-PRGM 251

RESULT 7

A72702

hypothetical protein APE1032 - Aeropyrum pernix (strain K1)

C:Species: Aeropyrum pernix

C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jun-2000

C:Accession: A72702

R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; T

awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J

DNA Res. 6, 83-101, 1999

A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aer

A:Reference number: A72450; MUID:99310339

A:Accession: A72702

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-363 <KAW>

A:Cross-references: DDBJ:AP000060; NID:g5104188; PIDN:BAAB0017.1; PID:d1043803; PID:

A:Experimental source: strain K1

C:Genetics:

A:Gene: APE1032

C:Superfamily: Aeropyrum pernix hypothetical protein APE1032

Query Match

Best Local Similarity

Matches 23; Conservative

18.5%; Score 62.5; DB 2; Length 363;

36.5%; Pred. No. 16;

Mismatches 20; Indels 15; Gaps 3;

OY

2

GDHKLARPAALQ-----DPLQSLRPRGHAPRRPRDLGGSGVQSDGVLOHLQRRGH 56

DB

78

GEYQLGVYPRLOYAVPVLEVPFSLRPSLRPPRRHDLRYALD-----VLQH---PAD 127

OY

57

VKL 59

DB

128

VGL 130

RESULT 8

S72211

N-acetyl-beta-D-glucosaminidase precursor - fungus (Trichoderma harzianum) (strain F

C:Species: Trichoderma harzianum

A:Variety: strain P1

C:Date: 19-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 21-Jan-2000

C:Accession: S72211

R:Peterbauer, C.K.; Lorito, M.; Hayes, C.K.; Harman, G.E.; Kubicek, C.P.

Curr. Genet. 30, 325-331, 1996

A:Title: Molecular cloning and expression of the nagI gene (N-acetyl-beta-D- glucosa

A:Reference number: S72211; MUID:96376496

A:Accession: S72211

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-580 <PEP>

A:Cross-references: GB:S83231; NID:g1911744; PIDN:AA050829.1; PID:g1911745

C:Superfamily: beta-hexosaminidase

Query Match

Best Local Similarity

Matches 19; Conservative

18.3%; Score 62; DB 2; Length 580;

37.3%; Pred. No. 31;

Mismatches 15; Indels 12; Gaps 4;

Oy 15 PRLQSLRPRGHAPHP-----PPRDLG-----GLDSGVQSDGVLOHLQRPCHV 57
 DB 237 PSLPRKLSQSG-AVHPSLVYSPADLAGIFQYGLDRGVE---VITEIDMPCHI 283

RESULT 9

A1145

gamma-aminobutyric acid receptor A, cyclodiene resistance-conferring - fruit fly (Drosophila)
 C:Species: Drosophila melanogaster
 C:Date: 10-Apr-1992 #sequence_revision 10-Apr-1992 #text_change 21-Aug-1998

C:Accession: A1145
 R:Refenc-Constant, R.H.; Mortlock, D.P.; Shaffer, C.D.; MacIntyre, R.J.; Roush, R.T.
 Proc. Natl. Acad. Sci. U.S.A. 88, 7209-7213, 1991

A:Title: Molecular cloning and transformation of cyclodiene resistance in Drosophila: an
 A:Reference number: A1145; MUID:91334435

A:Accession: A1145

A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
 A:Molecule type: mRNA

A:Residues: 1-606 <FFR>
 A:Cross-references: GB:M69057

C:Genetics:
 A:Gene: FlyBase:Rdl

A:Cross-references: FlyBase:FBgn0004244
 C:Superfamily: acetylcholine receptor

C:Keywords: neurotransmitter receptor; transmembrane protein

Query Match 18.3%; Score 62; DB 2; Length 606;
 Best Local Similarity 36.7%; Pred. No. 33;

Matches 18; Conservative 2; Mismatches 17; Indels 12; Gaps 2;

Oy 22 PRGHAPHP-----PPRDLGGLDSGVQSDG-----VLQHLQRPCHV 58
 DB 504 PQSHGPRPGGGGPGGGGGGGGPRGCGDPEAAVPAHLHLRGKVK 552

RESULT 10
 B72532
 hypothetical protein APE2230 - Aeropyrum pernix (strain K1)

C:Species: Aeropyrum pernix
 C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999

C:Accession: B72532
 R:Kawabuchi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Halkawa, Y.; Jin-no, K.; Takah

awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.; K
 DNA Res. 6, 83-101, 1999

A:Title: Complete genome sequence of an aerobic hyper-thermophilic Cremonarchaeon, Aeropyz
 A:Reference number: A72450; MUID:99310339

A:Accession: B72532
 A:Status: preliminary

A:Molecule type: DNA
 A:Residues: 1-266 <KAM>

A:Cross-references: DDBJ:AP000063; NID:g5105654; PIDN:BAA81242.1; PID:d1045028; PID:g510
 A:Experimental source: strain K1

C:Genetics:
 A:Gene: APE2230

Query Match 18.0%; Score 61; DB 2; Length 266;
 Best Local Similarity 28.6%; Pred. No. 17;

Matches 20; Conservative 10; Mismatches 16; Indels 24; Gaps 4;

Oy 8 HRPALQDPL---OSLRG-----HAPHPRL-----GGGLDSGVQ 44
 DB 106 HTPALEGVVLVYRAVAPQGRGVPLVLYLHRTGPRDLPCHLPRGAGLD-GLYE 164

Oy 45 DGVLOHLQRP 54
 DB 165 DGVDEGVDP 174

RESULT 11
 A96553

probable myrosinase precursor 53323-50499 (imported) - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
 C:Accession: A96553

R:Rheoelion, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,

ansen, N.F.; Hughes, B.; Hulzar, L.
 Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Malt, R.; Marzia

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719

A:Accession: A96553
 A:Status: preliminary

A:Molecule type: DNA
 A:Residues: 1-465 <STO>
 A:Cross-references: GB:AE005173; NID:g10092358; PIDN:AMG12767.1; GSPDB:GN00141
 C:Genetics:
 A:Gene: F5D21.17
 A:Map position: 1
 C:Superfamily: Agrobacterium beta-glucosidase

Query Match 17.9%; Score 60.5; DB 2; Length 465;
 Best Local Similarity 39.2%; Pred. No. 36;

Matches 20; Conservative 6; Mismatches 18; Indels 7; Gaps 3;

Oy 12 LQDPLQSLRPRGHAPHP-----DLGGGLDSGVQSDGVLOHLQRPCHV 58
 DB 164 LNPQSLALRKGNGSYPRGCTGCLGD--DSGVPEYTV-AHNDLHAHAK 211

RESULT 12
 S37671
 MHC class III histocompatibility antigen HLA-B-associated protein 2 [similarity] - hu

C:Species: Homo sapiens (man)
 C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 15-Sep-2000

C:Accession: S37671
 R:Bougueleret, L.

submitted to the EMBL Data Library, August 1992

A:Reference number: S37671

A:Accession: S37671

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1870 <BOU>

A:Cross-references: EMBL:Z15025; NID:g29374; PID:g29375

C:Genetics:
 A:Map position: 6p21.3

A:Introns: 38/2; 129/3; 154/1; 202/1; 252/3; 279/2; 327/1; 357/2; 429/3; 588/1;
 C:Superfamily: collagen alpha 1(IV) chain

Query Match 17.9%; Score 60.5; DB 2; Length 1870;
 Best Local Similarity 35.9%; Pred. No. 1.7e+02;

Matches 14; Conservative 7; Mismatches 15; Indels 3; Gaps 1;

Oy 22 PRGHAPHPRDLGGGLDSGVQSDGVLOHLQRPCHV 60
 DB 1502 PQGSPRPRTREPRQVNSGLSSD---PHEPRGMVNG 1537

RESULT 13
 S36152

MHC class III histocompatibility antigen HLA-B-associated protein 2 [similarity] - hu

C:Species: Homo sapiens (man)
 C:Date: 06-Jun-1995 #sequence_revision 17-Nov-1995 #text_change 15-Sep-2000

C:Accession: S36152
 R:Iris, F.J.M.; Bougueleret, L.; Prieur, S.; Caterina, D.; Primas, G.; Perrot, V.; Ju

Nature Genet. 3, 137-145, 1993
 A:Title: Dense Alu clustering and a potential new member of the NFkapab family with1
 A:Reference number: S36152; MUID:93272029

A:Accession: S36152
 A:Status: preliminary
 A:Molecule type: DNA

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2002, 08:56:03 ; Search time 10 Seconds

(without alignments)
240,061 Million cell updates/sec

Title: US-09-640-636A-4

Perfect score: 338
Sequence: 1 RGDHKLHRAHPALDPLQLSL.....QSDGVLOHLQRPCHVKLGTA 62

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	74.5	22.0	700	1 UVRD_MYCTU	053344 mycobacteri
2	66	19.5	500	1 CP46_HUMAN	096462 homo sapien
3	66	19.5	500	1 CP46_MOUSE	096462 mus musculu
4	64	18.9	146	1 CYTC_HUMAN	P01034 homo sapien
5	63	18.6	146	1 CYTC_MACMU	019092 macaca mula
6	62	18.3	606	1 GAB_DROME	P25123 drosophila
7	61.5	18.2	335	1 PLS1_RAT	P58195 rattus norv
8	61	18.0	320	1 YFHA_BORPE	P53445 bordetella
9	61	18.0	809	1 CMAD_HUMAN	008449 homo sapien
10	60.5	17.9	991	1 GBR1_RAT	092004 rattus norv
11	60.5	17.9	1258	1 GLI2_HUMAN	P10070 homo sapien
12	60.5	17.9	1242	1 BAT2_HUMAN	P48634 homo sapien
13	59	17.5	934	1 TOP1_MYCTU	059567 mycobacteri
14	59	17.5	1355	1 SALM_DROME	P39770 drosophila
15	58.5	17.3	146	1 CYTC_SAISC	O19093 salmistr sci
16	58.5	17.3	245	1 HXB4_CHICK	P14840 gallus galli
17	58.5	17.3	303	1 VLYB_BACSU	045480 bacillus su
18	58.5	17.3	303	1 GBR1_MOUSE	096462 mus musculu
19	58	17.2	382	1 MEI3_HUMAN	096462 homo sapien
20	58	17.2	525	1 HRG_HUMAN	P04196 homo sapien
21	58	17.2	646	1 HS70_THEAN	P16019 theileria a
22	57.5	17.0	325	1 OTP_MOUSE	009113 mus musculu
23	57.5	17.0	545	1 DCFP_AZOBR	P51852 azospirillu
24	57	16.9	419	1 AROA_METHH	O26860 methanobact
25	57	16.9	528	1 CYR1_YEAST	000957 saccharomyc
26	57	16.9	691	1 TOP1_BACSU	P39814 bacillus su
27	57	16.9	819	1 FYB_MOUSE	035601 mus musculu
28	57	16.9	864	1 YC18_HUMAN	096462 homo sapien
29	57	16.9	911	1 MAN1_HUMAN	096462 homo sapien
30	56.5	16.7	396	1 FLO_ANTYKA	P13463 agrobacteri
31	56.5	16.7	678	1 VID3_AGRRA	P14283 bordetella
32	56.5	16.7	910	1 PERT_BORPE	096462 homo sapien
33	56.5	16.7	961	1 GBR1_HUMAN	096462 homo sapien

34	56	16.6	154	1 Y05B_CAEEL	009254 caenorhabdi
35	56	16.6	298	1 34KD_MYCPA	004959 mycobacteri
36	56	16.6	299	1 PAF_CHICK	060415 citreellus
37	56	16.6	545	1 AIRE_HUMAN	043918 homo sapien
38	56	16.6	648	1 Y084_HUMAN	014699 homo sapien
39	56	16.6	823	1 YRR2_CAEEL	009345 caenorhabdi
40	56	16.6	947	1 TOP1_MYCLE	069548 mycobacteri
41	56	16.6	1369	1 MOKE_SCHRO	095704 schizosacch
42	55.5	16.4	245	1 VE4_HPV05	P06924 human papil
43	55.5	16.4	328	1 CEBB_CHICK	005826 gallus galli
44	55.5	16.4	649	1 FPC1_YEAST	P32367 saccharomyc
45	55.5	16.4	928	1 DPOL_SALTY	096173 salmonella

ALIGNMENTS

RESULT 1
ID UVRD_MYCTU STANDARD: PRT: 700 AA.
AC 053344;
DR 30-MAY-2000 (Rel. 39, Last sequence update)
DR 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable DNA helicase II homolog (EC 3.6.1.-).
GN UVRD OR RVJ3198C OR MT3291 OR MTV014.42C.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=96295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hovnsby T., Jagers K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrett B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
complete genome sequence.";
RL Nature 393:537-544(1998).
[2]
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A.L., Uitterlind T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
laboratory strains.";
RL Submitted (Apr 2001) to the EMBL/Genbank/DBJ databases.
-i- FUNCTION: HAS BOTH ATPASE AND HELICASE ACTIVITIES. UNWINDS DNA
DUPLICATES WITH 3' TO 5' POLARITY WITH RESPECT TO THE BOUND STRAND
AND INITIATES UNWINDING MOST EFFECTIVELY WHEN A SINGLE-STRANDED
REGION IS PRESENT. INVOLVED IN THE POSTINSCRIPTION EVENTS OF
NUCLEOTIDE EXCISION REPAIR AND METHYL-DIRECTED MISMATCH REPAIR
(BY SIMILARITY).
-i- SIMILARITY: BELONGS TO THE HELICASE FAMILY. UVRD SUBFAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).

EMBL: AL021646; CNA16663.1; -

```

DR EMBL: AE007141; AAA47634.1; -.
DR HSSP: P36255; 1PDR.
DR TIGR: MT3291; -.
DR Tuberculist: RV3198C; -.
DR InterPro: IPR002121; HRDC.
DR InterPro: IPR000212; UVRD-helicase.
DR Pfam: PF00570; HRDC; 1.
DR Pfam: PF00580; UVRD-helicase; 1.
DR SMART: SM00341; HRDC; 1.
DR DNA repair: DNA replication; Helicase; ATP-binding; DNA-binding;
KW Hydrolyase; Complete proteome.
FT NP_BIND 31
SQ SEQUENCE 700 AA; 75603 MW; 60A1B7520855C81C CRC64;

Query Match 22.0%; Score 74.5; DB 1; Length 700;
Best Local Similarity 62.1%; Pred. No. 0.69;
Matches 18; Conservative 1; Mismatches 9; Indels 1; Gaps 1;

OY 3 DHKLAHPALDP-LLOSLRPRGHAPPP 30
Db 463 DDELAQRPALDLPGLLAELRRRAREARHPP 491

RESULT 2
CP46_HUMAN STANDARD; PRT; 500 AA.
AC 09Y6A2;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Cytochrome p450 46 (EC 1.14.--) (Cholesterol 24-hydroxylase).
GN CYP46.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal brain;
RX MEDLINE=99307393; PubMed=10377398;
RA Lund E.G., Guilevardo J.M., Russell D.W.;
RT "cDNA cloning of cholesterol 24-hydroxylase, a mediator of cholesterol
homeostasis in the brain.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:7238-7243(1999).
CC -1- FUNCTION: INVOLVED IN THE TURNOVER OF CHOLESTEROL. IT CONVERTS
CHOLESTEROL INTO 24S-HYDROXYCHOLESTEROL AND, TO A LESSER EXTENT,
25-HYDROXYCHOLESTEROL.
CC -1- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM MEMBRANE. MAY BE
ANCHORED TO THE MEMBRANE VIA A SINGLE TRANSMEMBRANE DOMAIN.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN BRAIN. MRNA WAS BROADLY
DISTRIBUTED WITH HIGHER LEVELS IN GRAY MATTER ZONES AND LOWER
LEVELS IN REGIONS RICH IN WHITE MATTER. NOT DETECTED IN FETAL
SAMPLE BUT ITS EXPRESSION INCREASES LINEARLY WITH AGE.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL: AF094480; AAD41244.1; -.
DR MGI: 604087; -.
DR InterPro: IPR001128; Cyt_P450.
DR Pfam: PF00067; P450; 1.
DR PRINTS: PR00385; P450.
DR PROSITE: PS00086; CYTOCHROME_P450; 1.
KW Oxidoreductase; Monooxygenase; Electron transport; Transmembrane;
KW Heme; Microsome; Endoplasmic reticulum.
FT TRANSMEM 3
FT DOMAIN 493 499 POLY-PRO.
* DOMAIN 493 499 POLY-PRO.

```

```

FT BINDING 437 437 HEME (BY SIMILARITY).
SQ SEQUENCE 500 AA; 56821 MW; AB9307749D9E5FDA CRC64;

Query Match 19.5%; Score 66; DB 1; Length 500;
Best Local Similarity 66.7%; Pred. No. 4.2;
Matches 12; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 14 DPLQSLRPRGHAPPP 31
Db 480 DPLVCTLRPRGQAPPP 497

RESULT 3
CP46_MOUSE STANDARD; PRT; 500 AA.
ID CP46_MOUSE
AC 09WYK8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Cytochrome p450 46 (EC 1.14.--) (Cholesterol 24-hydroxylase).
GN CYP46.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=99307393; PubMed=10377398;
RA Lund E.G., Guilevardo J.M., Russell D.W.;
RT "cDNA cloning of cholesterol 24-hydroxylase, a mediator of cholesterol
homeostasis in the brain.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:7238-7243(1999).
CC -1- FUNCTION: INVOLVED IN THE TURNOVER OF CHOLESTEROL. IT CONVERTS
CHOLESTEROL INTO 24S-HYDROXYCHOLESTEROL AND, TO A LESSER EXTENT,
25-HYDROXYCHOLESTEROL.
CC -1- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM MEMBRANE. MAY BE
ANCHORED TO THE MEMBRANE VIA A SINGLE TRANSMEMBRANE DOMAIN.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN HIGH LEVEL IN BRAIN (CEREBRAL
CORTEX, HIPPOCAMPUS, DENTATE GYRUS AND THALAMUS) AND IN LOWER
LEVELS IN TESTIS AND LIVER. BRAIN'S EXPRESSION INCREASED LINEARLY
WITH AGE.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL: AF094479; AAD41243.1; -.
DR MGI: MGI:1341877; CYP46.
DR InterPro: IPR001128; Cyt_P450.
DR Pfam: PF00067; P450; 1.
DR PRINTS: PR00385; P450.
DR PROSITE: PS00086; CYTOCHROME_P450; 1.
KW Oxidoreductase; Monooxygenase; Electron transport; Transmembrane;
KW Heme; Microsome; Endoplasmic reticulum.
FT TRANSMEM 3
FT DOMAIN 493 499 POLY-PRO.
FT BINDING 437 437 HEME (BY SIMILARITY).
SQ SEQUENCE 500 AA; 56814 MW; POF17BF65820F2 CRC64;

Query Match 19.5%; Score 66; DB 1; Length 500;
Best Local Similarity 66.7%; Pred. No. 4.2;
Matches 12; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 14 DPLQSLRPRGHAPPP 31
Db 480 DPLVCTLRPRGQAPPP 497

```

RESULT 4
CYTC_HUMAN
ID CYTC_HUMAN STANDARD: PRT: 146 AA.
AC P01034;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Cystatin C precursor (Neuroendocrine basic polypeptide) (Gamma-trace)
DE (Post-gamma globulin).
GN CST3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Placenta;
RX MEDLINE=87219149; PubMed=3495457;
RA Abrahamson M., Grubb A., Olafsson I., Lundwall A.;
RT "Molecular cloning and sequence analysis of cDNA coding for the
RL precursor of the human cysteine proteinase inhibitor cystatin C.";
FEBs Lett. 216:229-233(1987).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-Leukocyte;
RX MEDLINE=90303202; PubMed=2363674;
RA Abrahamson M., Olafsson I., Palsdottir A., Ulsbaeck M., Lundwall A.,
RT Jenson O., Grubb A.;
RL "Structure and expression of the human cystatin C gene.";
Biochem. J. 268:287-294(1990).
RN [3]
RP SEQUENCE FROM N.A. (HCHWA VARIANT).
RC TISSUE-Brain;
RX MEDLINE=89235594; PubMed=2541223;
RA Levy E., Lopez-Otin C., Ghiso J., Gellner D., Frangione B.;
RT "Stroke in Icelandic patients with hereditary amyloid angiopathy is
RL related to a mutation in the cystatin C gene, an inhibitor of
cysteine proteases.";
J. Exp. Med. 169:1771-1778(1989).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=89350949; PubMed=2764935;
RA Saito E., Sabatini L.M., Eddy R.L., Shows T.B., Azen E.A.,
RT Isemura S., Sanada K.;
RL "The human cystatin C gene (CST3) is a member of the cystatin gene
family which is localized on chromosome 20.";
Biochem. Biophys. Res. Commun. 162:1324-1331(1989).
RN [5]
RP SEQUENCE FROM N.A.
RA Dickinson D.P., Hewett-Emmett D., Thiesse M.;
RT "Acquisition of complex patterns of differential expression in
RT epithelial cell populations during the evolution of type 2 cystatin
genes.";
Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RA Delours P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RT Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagquley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA Buck D., Burrill W., Butler A.P., Carder C., Carter N.P.,
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA Clegg S., Cobley V.E., Collier R.E., Connor C., Corby N.R.,
RA Coulson A., Coville G.J., Deadman R., Dhani P., Dunn M.,
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA Graffam D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA Levaialaho M.H., Levesha M., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McConachie L.J., McMay K., McMurray A.A.,
RA Milne S., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,

RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Philimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Shonkhen R., Sims S.,
RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
RA Swann M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.;
RT "The DNA sequence and comparative analysis of human chromosome 20.";
Nature 414:865-871(2001).
RN [7]
RP SEQUENCE FROM N.A.
RC TISSUE-Brain;
RA Strausberg R.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
RN [8]
RP SEQUENCE OF 27-146.
RX MEDLINE=82222268; PubMed=6283552;
RA Grubb A., Loebberg H.;
RT "Human gamma-trace, a basic microprotein: amino acid sequence and
RT presence in the adenohiphysis.";
Proc. Natl. Acad. Sci. U.S.A. 79:3024-3027(1982).
RN [9]
RP SEQUENCE OF 27-73.
RX MEDLINE=84110059; PubMed=6662498;
RA Turk V., Brzin J., Longer M., Ritonja A., Eropkin M., Borchart U.,
RA Macleleld W.;
RT "Protein inhibitors of cysteine proteinases. III. Amino-acid sequence
RT of cystatin from chicken egg white.";
Hoppe-Seyler's Z. Physiol. Chem. 364:1487-1496(1983).
RN [10]
RP SEQUENCE OF 27-76.
RX MEDLINE=84128015; PubMed=6365094;
RA Brzin J., Popovic T., Turk V.;
RT "Human cystatin, a new protein inhibitor of cysteine proteinases.";
Biochem. Biophys. Res. Commun. 118:103-109(1984).
RN [11]
RP DISULFIDE BONDS.
RA Grubb A., Loebberg H., Barrett A.J.;
RT "The disulphide bridges of human cystatin C (gamma-trace) and chicken
RT cystatin.";
FEBs Lett. 170:370-374(1984).
RN [12]
RP X-RAY CRYSTALLOGRAPHY (2.50 ANGSTROMS) OF 27-146.
RX MEDLINE=21173909; PubMed=11276250;
RA Janowski R., Kozak M., Jankowska E., Grzonka Z., Grubb A.,
RA Abrahamson M., Jaskolski M.;
RT "Human cystatin C, an amyloidogenic protein, dimerizes through
RT three-dimensional domain swapping.";
Nat. Struct. Biol. 8:316-320(2001).
RN [13]
RP VARIANT GLN-94.
RX MEDLINE=92316504; PubMed=1352269;
RA Abrahamson M., Jonsdottir S., Olafsson I., Jenson O., Grubb A.;
RT "Hereditary cystatin C amyloid angiopathy: identification of the
RT disease-causing mutation and specific diagnosis by polymerase chain
RT reaction based analysis.";
Hum. Genet. 89:377-380(1992).
RN [14]
RP FUNCTION: AS AN INHIBITOR OF CYSTEINE PROTEINASES, THIS PROTEIN IS
THOUGHT TO SERVE AN IMPORTANT PHYSIOLOGICAL ROLE AS A LOCAL
REGULATOR OF THIS ENZYME ACTIVITY.
CC -1- SUBUNIT: HOMODIMER.
CC -1- TISSUE SPECIFICITY: IT IS EXPRESSED IN HIGHEST LEVELS IN THE
EPIDIDYMIS, VAS DEFERENS, BRAIN, THYMUS, AND OVARY AND THE LOWEST
IN THE SUBMANDIBULAR GLAND.
CC -1- DISEASE: IMPLICATED IN A HEREDITARY FORM OF CEREBRAL HEMORRHAGE
CHARACTERIZED BY A THICKENING OF THE CEREBRAL ARTERIES WALLS WITH
DEPOSITION OF MATERIAL WITH THE CHARACTERISTICS OF AMYLOID. THIS
DISEASE IS KNOWN AS HEREDITARY CEREBRAL HEMORRHAGE WITH
AMYLOIDOSIS (HCHWA), OR HEREDITARY CYSTATIN C AMYLOID ANGIOPATHY
(HCCAA) OR ALSO AS ICELANDIC CEREBROVASCULAR TYPE AMYLOIDOSIS.
CC -1- SIMILARITY: BELONGS TO THE CYSTATIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; X05607; CAA29096.1; -;
DR EMBL; X52255; CAA36497.1; -;
DR EMBL; M27891; AAA52164.1; -;
DR EMBL; M27889; AAA52164.1; JOINED.
DR EMBL; M27890; AAA52164.1; JOINED.
DR EMBL; X61681; CAA43856.2; -;
DR EMBL; X61682; CAA43856.2; JOINED.
DR EMBL; X61683; CAA43856.2; JOINED.
DR EMBL; AF319564; AAK11570.1; -;
DR EMBL; AL121894; CAC05424.1; -;
DR EMBL; BC013083; AAI13083.1; -;
DR PIR; S00004; UDHO.
DR PIR; J10095; J10095.
DR PIR; A32732; A32732.
DR PIR; A33400; A33400.
DR PIR; S10216; S10216.
DR PDB; 1G96; 06-APR-01.
DR MIM; 604312; -;
DR MIM; 105150; -;
DR InterPro; IPR000010; Cystatin.
DR InterPro; IPR003243; Cystatin_C_M.
DR Pfam; PF00031; cystatin; 1.
DR ProDom; PD001231; Cystatin_C_M; 1.
DR SMART; SM00043; CY; 1.
DR PROSITE; PS00287; CYSTATIN; 1.
KW Thiol protease inhibitor; Amyloid; Signal; Disease mutation;
KW Polymorphism; 3D-structure.
FT SIGNAL 1 26
FT CHAIN 27 146 CYSTATIN C.
FT ACT_SITE 37 37 REACTIVE SITE.
FT SITE 81 85 SECONDARY AREA OF CONTACT.
FT DISULFID 99 109
FT DISULFID 123 143
FT VARIANT 25 25 A -> T (IN DBSNP:1064039).
FT VARIANT 94 94 /FTID=VAR.011893.
FT VARIANT L -> Q (IN HCHMA).
FT VARIANT /FTID=VAR.002207.
FT CONFLICT 132 132 W -> S (IN REF.3).
FT SEQUENCE 146 AA; 15799 MW; 75EF049CAE2E8B2B CRC64;
Query Match 18.9%; Score 64; DB 1; Length 146;
Best Local Similarity 42.9%; Pred. No. 1.9;
Matches 12; Conservative 7; Mismatches 9; Indels 0; Gaps 0;
OY 24 GHAPPPPRDLGGDGSQSDGYLQHL 51
DB 26 GSSPGKPPRLVGGPMDASVEEGVRRAL 53
RESULT 5
CYTC_MACMU
ID CYTC_MACMU STANDARD; PRT; 146 AA.
AC 019092;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Cystatin C precursor.
GN CST3.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9544;
RN [1]

RP SEQUENCE FROM N.A.
RX MEDLINE=97054523; PubMed=8898820;
RA Wei L.H., Walker L.C., Levy E.;
RT "Cystatin C, Icelandic-like mutation in an animal model of
RT cerebrovascular beta-amyloidosis.";
RL Stroke 27:2080-2085(1996).
CC -I- FUNCTION: AS AN INHIBITOR OF CYSTEINE PROTEINASES, THIS PROTEIN IS
CC THOUGHT TO SERVE AN IMPORTANT PHYSIOLOGICAL ROLE AS A LOCAL
CC REGULATOR OF THIS ENZYME ACTIVITY.
CC -I- SIMILARITY: BELONGS TO THE CYSTATIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; U51912; AAB64050.1; -;
DR HSSP; P01038; IAB90.
DR InterPro; IPR000010; Cystatin.
DR InterPro; IPR003243; Cystatin_C_M.
DR Pfam; PF00031; cystatin; 1.
DR ProDom; PD001231; Cystatin_C_M; 1.
DR SMART; SM00043; CY; 1.
DR PROSITE; PS00287; CYSTATIN; 1.
KW Thiol protease inhibitor; Amyloid; Signal;
KW Polymorphism; 3D-structure.
FT SIGNAL 1 26
FT CHAIN 27 146 CYSTATIN C.
FT ACT_SITE 37 37 REACTIVE SITE.
FT SITE 81 85 SECONDARY AREA OF CONTACT.
FT DISULFID 99 109 BY SIMILARITY.
FT DISULFID 123 143 BY SIMILARITY.
FT SEQUENCE 146 AA; 15857 MW; F0B3BB774A29DF26 CRC64;
Query Match 18.6%; Score 63; DB 1; Length 146;
Best Local Similarity 42.9%; Pred. No. 2.4;
Matches 12; Conservative 7; Mismatches 9; Indels 0; Gaps 0;
OY 24 GHAPPPPRDLGGDGSQSDGYLQHL 51
DB 26 GSSPGKPPRLVGGPMDASVEEGVRRAL 53
RESULT 6
GAB_DROME
ID GAB_DROME STANDARD; PRT; 606 AA.
AC P25123; Q26302;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Gamma-aminobutyric-acid receptor beta subunit precursor (GABA(A)
DE receptor) (Cyclodiene resistance protein).
GN RDL.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=91334435; PubMed=1651498;
RA Ffrench-Constant R.H., Mortlock D.P., Shaffer C.D.,
RA Macintyre R.J., Roush R.T.;
RT "Molecular cloning and transformation of cyclodiene resistance in
RT Drosophila: an invertebrate gamma-aminobutyric acid subtype A
RT receptor locus.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:7209-7213(1991).
RN [2]
SEQUENCE OF 70-113 FROM N.A.
RX MEDLINE=93260477; PubMed=7684073;
RA Ffrench-Constant R.H., Rochelleau T.A.;


```

RT "Drosophila gamma-aminobutyric acid receptor gene Rdl shows extensive
RT alternative splicing.":
RL J. Neurochem. 60:2323-2326(1993).
CC -1- FUNCTION: GABA, AN INHIBITORY NEUROTRANSMITTER, MEDIATES NEURONAL
CC INHIBITION BY BINDING TO THE GABA RECEPTOR AND OPENING AN INTEGRAL
CC CHLORIDE CHANNEL.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- MISCELLANEOUS: RESISTANCE IS THOUGHT TO BE DUE TO INSENSITIVITY OF
CC THE CYCLODIENE/PICROTOXININ BINDING SITE ON THE GABA(A) RECEPTOR-
CC CHLORIDE IONOPHORE COMPLEX.
CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.
CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1, MET-6 OR MET-12 IS THE
CC INITIATOR.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: M69057; AAA28556.1; -.
CC EMBL: S61113; AAB26669.1; -.
CC PIR: A41145; A41145.
CC FLYBASE: FBgn0004244; Rdl.
CC InterPro: IPR000188; GABA_A_receptor.
CC InterPro: IPR001175; Neur_channel.
CC Pfam: PF02931; Neur_chan_LBD; 1.
CC Pfam: PF02932; Neur_chan_LBD; 1.
CC PRINTS: PR00252; NRIONCHANNEL.
CC PROSITE: PS00236; NEUROTR_ION_CHANNEL; 1.
CC Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;
CC Transmembrane.
CC FT SIGNAL 1 44 POTENTIAL.
CC FT CHAIN 45 606 GAMMA-AMINOBUTYRIC-ACID RECEPTOR BETA
CC SUBUNIT.
CC FT DOMAIN 45 268 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 269 291 POTENTIAL.
CC FT TRANSMEM 297 316 POTENTIAL.
CC FT TRANSMEM 333 356 POTENTIAL.
CC FT DOMAIN 357 568 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 569 590 POTENTIAL.
CC FT CARBOHYD 58 58 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 248 248 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 253 253 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT DISULFID 185 199 BY SIMILARITY.
CC FT DOMAIN 499 535 GLY-RICH.
CC FT CONFLICT 97 97 V -> L (IN REF. 2).
CC FT CONFLICT 101 101 L -> K (IN REF. 2).
CC FT SEQUENCE 606 AA; 65748 MW; 81B9DB08E1906EF1 CRC64;

Query Match 18.3%; Score 62; DB 1; Length 606;
Best Local Similarity 36.7%; Pred. No. 14;
Matches 18; Conservative 2; Mismatches 17; Indels 12; Gaps 2;

OY 22 PRGAPHP-----PPRDLGGGIDSGVQSDG-----VLQHLQRPQHYK 58
Db 504 PQSHQPGGQGGGPGGGGGGGGGGGGDPPEAAVPAHLHLPKVK 552

RESULT 7
PLS1_RAT STANDARD; PRT; 335 AA.
AC P58195;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Phospholipid scramblase 1 (PL scramblase 1) (Ca2+ dependent
DE Phospholipid scramblase 1).
GN PLSCR1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

```

```

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND PHOSPHORYLATION.
RX MEDLINE=21282879; PubMed=11259432;
RA Pastorelli C., Veiga J., Charles N., Volgnier E., Moussu H.,
RA Monteiro R.C., Benhamou M.;
RT "Ige receptor type I-dependent tyrosine phosphorylation of
RT phospholipid scramblase.";
RL J. Biol. Chem. 276:20407-20412(2001).
CC -1- FUNCTION: MAY MEDiate ACCELERATED ATP-INDEPENDENT BIDIRECTIONAL
CC TRANSLAYER MIGRATION OF PHOSPHOLIPIDS UPON BINDING CALCIUM IONS
CC THAT RESULTS IN A LOSS OF PHOSPHOLIPID ASYMMETRY IN THE PLASMA
CC MEMBRANE. MAY PLAY A CENTRAL ROLE IN THE INITIATION OF FIBRIN CLOT
CC FORMATION, IN THE ACTIVATION OF MAST CELLS AND IN THE RECOGNITION
CC OF APOPTOTIC AND INJURED CELLS BY THE RETICULOENDOTHELIAL SYSTEM.
CC -1- CORFACTOR: CALCIUM ION (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Type II membrane protein (by similarity).
CC -1- PTM: PHOSPHORYLATED ON TYROSINE RESIDUES.
CC -1- SIMILARITY: BELONGS TO THE PHOSPHOLIPID SCRAMBLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: AY024347; AAK00575.1; -.
CC Transmembrane; Lipoprotein; Calcium-binding; SH3-binding; Repeat;
CC Phosphorylation; Palmitate.
CC FT DOMAIN 1 305 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 306 322 POTENTIAL.
CC FT TRANSMEM 323 335 EXTRACELLULAR (POTENTIAL).
CC FT DOMAIN 323 335 7 X 7 AA TANDEM REPEATS OF Q-G-P-Y-[AP]-
CC DOMAIN 23 71 G-P.
CC FT REPEAT 23 29 1.
CC FT REPEAT 30 36 2.
CC FT REPEAT 37 43 3.
CC FT REPEAT 44 50 4.
CC FT REPEAT 51 57 5.
CC FT REPEAT 58 64 6.
CC FT REPEAT 65 71 7 (APPROXIMATE).
CC FT SITE 64 72 SH3-BINDING 1 (POTENTIAL).
CC FT SITE 101 109 SH3-BINDING 2 (POTENTIAL).
CC FT DOMAIN 198 206 CYS-RICH.
CC FT DOMAIN 106 109 POLY-PRO.
CC FT MOD_RES 91 91 PHOSPHORYLATION (BY ABL) (BY SIMILARITY).
CC FT MOD_RES 178 178 PHOSPHORYLATION (BY PKC) (BY SIMILARITY).
CC FT LIPID 251 251 PALMITATE (PROBABLE).
CC FT LIPID 254 254 PALMITATE (PROBABLE).
CC FT LIPID 256 256 PALMITATE (PROBABLE).
CC FT LIPID 257 257 PALMITATE (PROBABLE).
CC SEQUENCE 335 AA; 36711 MW; 7A674DDEF2759EB CRC64;

Query Match 18.2%; Score 61.5; DB 1; Length 335;
Best Local Similarity 40.0%; Pred. No. 8.6;
Matches 18; Conservative 3; Mismatches 23; Indels 1; Gaps 1;

OY 8 HRPALDPLLSLRGRGHPHPPRDLGGGIDSGVQSDVQLQHLQ 52
Db 87 HQPAYNHPGPGGPGGTPWQAP-PPPLDCEPLGYLQIDQILVHQQ 130

RESULT 8
YFHA_BORPE STANDARD; PRT; 320 AA.
ID YFHA_BORPE
AC P33445;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 33.8 kDa protein in FHAC 3 region (ONFA).

```


Db 35 HE0HHQYPLRQPOFRLHPHNLPPPPP 62

RESULT 10

GBRL_RAT STANDARD: PRT: 991 AA.

AC 092004; 008620; 008621; 0920F9; 092308;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE gamma-aminobutyric acid type B receptor, subunit 1 precursor (GABA-B receptor 1) (GABA-B-R1) (Gbl1).

GN GABBR1.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxID=10116;

RA KAUPMANN K., HUGGEL K., HEID J., FLOR P.J., BISCHOFF S., MICKEL S.J., MACHMANN G., ANGST C., BITTIGER H., FROESTL W., BETTLER B.;

RA KAUPMANN K., HUGGEL K., HEID J., FLOR P.J., BISCHOFF S., MICKEL S.J., MACHMANN G., ANGST C., BITTIGER H., FROESTL W., BETTLER B.;

RT "Expression cloning of GABA(B) receptors uncovers similarity to metabotropic glutamate receptors.";

RL Nature 386:239-246(1997).

RN [2]

RP SEQUENCE FROM N.A. (ISOFORMS 1A AND 1B).

RC STRAIN-RICO: TISSUE-Brain cortex, and Cerebellum;

RX MEDLINE=97222131; PubMed=9069281;

RA ISOMOTO S., KALBARA M., SAKURAI-YAMASHITA Y., NAGAYAMA Y., UZONO Y., YANO K., TANIYAMA K.;

RT "Cloning and tissue distribution of novel splice variants of the rat GABAB receptor.";

RL Biochem. Biophys. Res. Commun. 233:10-15(1998).

RN [3]

RP SEQUENCE FROM N.A. (ISOFORM 1E).

RC STRAIN-Wistar: TISSUE-Hippocampus;

RX MEDLINE=99388283; PubMed=10457184;

RA PLATF T., MALITSCHKE B., KAUPMANN K., PREZEAU L., PIN J.-P., BETTLER B., KARSCHIN A.;

RT "Alternative splicing generates a novel isoform of the rat metabotropic GABA(B)R1 receptor.";

RL Eur. J. Neurosci. 11:2874-2882(1999).

RN [4]

RP SEQUENCE FROM N.A. (ISOFORM 1A), AND R1A-R2 INTERACTION.

RC TISSUE-Brain;

RX MEDLINE=99175124; PubMed=10075644;

RA NG G.Y.K., CLARK J., COULOMBE N., EHLER T.E., SULLIVAN R., KARGMAN S., CHATEAUNEUF A., TSUKUMOTO N., McDONALD T., WHITING P., MESEY E., JOHNSON M.P., LIU Q., KOLAKOWSKI L.F. JR., EVANS J.F.;

RA BONNET T.L., O'NEILL G.P.;

RT "Identification of a GABAB receptor subunit, gb2, required for functional GABAB receptor activity.";

RL J. Biol. Chem. 274:7607-7610(1999).

RN [5]

RP TISSUE DISTRIBUTION.

RX MEDLINE=20121644; PubMed=10658574;

RA BELLEY M., SULLIVAN R., REEVES A., EVANS J.F., O'NEILL G.P.;

RT "Synthesis of the nanomolar photoaffinity GABA(B) receptor ligand CGP 71872 reveals diversity in the tissue distribution of GABA(B) receptor forms.";

RL Bioorg. Med. Chem. 7:2697-2704(1999).

RN [6]

RP R1A-R2 INTERACTION.

RX MEDLINE=99102694; PubMed=9872744;

RA KUNER R., KOEHR G., GRUNEWALD S., EISENHARDT G., BACH A., KORNAU H.-C.;

RT "Role of heteromer formation in GABA-B Receptor Function.";

RL Science 283:74-77(1999).

RN [7]

RP MUTAGENESIS OF SER-247; SER-268 AND SER-269.

RX MEDLINE=20159055; PubMed=10692480;

RA GALVEZ T., URWYLER S., PREZEAU L., MOSBACHER J., JOLY C., MALITSCHKE B., HEID J., BRABET I., FROESTL W., BETTLER B., KAUPMANN K., PIN J.-P.;

RA KAUPMANN K., PIN J.-P.;

RT binding at GABA(B) receptors: involvement of serine 269 of the R1 GABA(B)R1 subunit.";

RL Mol. Pharmacol. 57:419-426(2000).

CC -1- FUNCTION: RECEPTOR FOR GABA. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G-PROTEINS THAT INHIBITS ADENYLYL CYCLASE ACTIVITY, STIMULATES PHOSPHOLIPASE A2, ACTIVATES POTASSIUM CHANNELS, INACTIVATES VOLTAGE-DEPENDENT CALCIUM-CHANNELS AND MODULATES INOSITOL PHOSPHOLIPIDS HYDROLYSIS. PLAYS A CRITICAL ROLE IN THE FINE-TUNING OF INHIBITORY SYNAPTIC TRANSMISSION. PRE-SYNAPTIC GABA-B-R INHIBIT NEUROTRANSMITTER RELEASE BY DOWN-REGULATING HIGH-VOLTAGE ACTIVATED CALCIUM CHANNELS, WHEREAS POSTSYNAPTIC GABA-B-R DECREASE NEURONAL EXCITABILITY BY ACTIVATING A PROMINENT INWARDLY RECTIFYING POTASSIUM (KIR) CONDUCTANCE THAT UNDERLIES THE LATE INHIBITORY POSTSYNAPTIC POTENTIALS. NOT ONLY IMPLICATED IN SYNAPTIC INHIBITION BUT ALSO IN HIPPOCAMPAL LONG-TERM POTENTIATION, SLOW WAVE SLEEP, MUSCLE RELAXATION AND ANTINOCICEPTION.

CC -1- COFACTOR: REQUIRES CALCIUM FOR HIGH AFFINITY BINDING TO GABA.

CC -1- SUBUNIT: HETERODIMER OF GABA-B-R1 AND GABA-B-R2. NEITHER OF WHICH IS EFFECTIVE ON ITS OWN AND HOMODIMERIC ASSEMBLY DOES NOT SEEM TO HAPPEN.

CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MOREOVER COEXPRESSION OF GABA-B-R1 AND GABA-B-R2 APPEARS TO BE A PREREQUISITE FOR MATURATION AND TRANSPORT OF GABA-B-R1 TO THE PLASMA MEMBRANE.

CC -1- ALTERNATIVE PRODUCTS: 5 ISOFORMS: 1A, 1B, 1C, 1D AND 1E (SHOWN HERE); ARE PRODUCED BY ALTERNATIVE SPLICING.

CC -1- TISSUE SPECIFICITY: ISOFORMS 1A, 1B AND 1C ARE EXPRESSED IN TESTIS, STOMACH, SPINAL CORD AND BRAIN INCLUDING CEREBRAL CORTICAL LAYERS, PYRAMIDAL CELL LAYERS OF THE HIPPOCAMPUS, GRANULAR CELL LAYERS OF THE DENTATE GYRUS, BASAL GANGLIA, CEREBELLUM (PREDOMINANTLY IN PURKINJE CELLS FOLLOWED BY GRANULAR LAYER).

CC -1- ISOFORM 1B IS ALSO EXPRESSED IN KIDNEY AND LIVER. ISOFORM 1D IS EXPRESSED IN FOREBRAIN, CEREBELLUM, EYE, KIDNEY, AND URINARY BLADDER.

CC -1- DOMAIN: ALPHA-HELICAL PARTS OF THE C-TERMINAL INTRACELLULAR REGION MEDIATE HETERODIMERIC INTERACTION WITH GABA-B RECEPTOR 2. THE LINKER REGION BETWEEN THE TRANSMEMBRANE DOMAIN 3 (TM 3) AND THE TRANSMEMBRANE DOMAIN 4 (TM 4) PROBABLY PLAY A ROLE IN THE SPECIFICITY FOR G-PROTEIN COUPLING.

CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.

CC -1- SIMILARITY: CONTAINS 2 SUSHI (SCR) DOMAINS.

CC -1- CAUTION: ISOFORM 1E HAS BEEN CALLED 1C IN REF. 3.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: Y10369; CAAT71398.1; -

DR EMBL: Y10370; CAAT71399.1; -

DR EMBL: AB016160; BAA14708.1; -

DR EMBL: AB016161; BAA14709.1; -

DR EMBL: AF110797; AAD19656.1; -

DR EMBL: AF110797; AAD19656.1; JOINED.

DR EMBL: AF110797; AAD19657.1; -

DR EMBL: AF110797; AAD19657.1; JOINED.

DR EMBL: AF110797; AAD19658.1; -

DR EMBL: AF110797; AAD19658.1; JOINED.

DR EMBL: AF110797; AAD19659.1; -

DR EMBL: AF110797; AAD19659.1; JOINED.

DR HSSP: P10998; 1YVD.

DR InterPro: IPR001828; ANF_Receptor.

DR InterPro: IPR000337; GPCR_Mgr.

```

DR InterPro: IPR000436; Sushi_SCR_CCP.
DR Pfam: PF00003; 7tm_3; 1.
DR Pfam: PF01094; ANF_receptor; 2.
DR Pfam: PF00084; sushi; 2.
DR PRINTS: PRO1176; GABARRECEPR.
DR PRINTS: PRO1177; GABABRECEPR.
DR SMART: SM00032; CCP; 2.
DR PROSITE: PS00979; G_PROTEIN_RECEPTOR_F3_1; FALSE_NEG.
DR PROSITE: PS00980; G_PROTEIN_RECEPTOR_F3_2; FALSE_NEG.
DR PROSITE: PS00981; G_PROTEIN_RECEPTOR_F3_3; FALSE_NEG.
DR PROSITE: PS00982; G_PROTEIN_RECEPTOR_F3_4; 1.
DR G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW Postsynaptic membrane; Coiled coil; Repeat; Alternative splicing.
FT SIGNAL 1 16
FT CHAIN 1 991
FT DOMAIN 17 590
FT TRANSFAM 591 611
FT TRANSFAM 612 630
FT TRANSFAM 631 651
FT TRANSFAM 652 666
FT TRANSFAM 667 687
FT TRANSFAM 688 709
FT TRANSFAM 710 730
FT TRANSFAM 731 757
FT TRANSFAM 758 798
FT TRANSFAM 819 834
FT TRANSFAM 835 855
FT TRANSFAM 856 863
FT TRANSFAM 864 884
FT TRANSFAM 884 904
FT TRANSFAM 904 924
FT TRANSFAM 924 944
FT TRANSFAM 944 964
FT TRANSFAM 964 984
FT TRANSFAM 984 1000
FT TRANSFAM 1000 1020
FT TRANSFAM 1020 1040
FT TRANSFAM 1040 1060
FT TRANSFAM 1060 1080
FT TRANSFAM 1080 1100
FT TRANSFAM 1100 1120
FT TRANSFAM 1120 1140
FT TRANSFAM 1140 1160
FT TRANSFAM 1160 1180
FT TRANSFAM 1180 1200
FT TRANSFAM 1200 1220
FT TRANSFAM 1220 1240
FT TRANSFAM 1240 1260
FT TRANSFAM 1260 1280
FT TRANSFAM 1280 1300
FT TRANSFAM 1300 1320
FT TRANSFAM 1320 1340
FT TRANSFAM 1340 1360
FT TRANSFAM 1360 1380
FT TRANSFAM 1380 1400
FT TRANSFAM 1400 1420
FT TRANSFAM 1420 1440
FT TRANSFAM 1440 1460
FT TRANSFAM 1460 1480
FT TRANSFAM 1480 1500
FT TRANSFAM 1500 1520
FT TRANSFAM 1520 1540
FT TRANSFAM 1540 1560
FT TRANSFAM 1560 1580
FT TRANSFAM 1580 1600
FT TRANSFAM 1600 1620
FT TRANSFAM 1620 1640
FT TRANSFAM 1640 1660
FT TRANSFAM 1660 1680
FT TRANSFAM 1680 1700
FT TRANSFAM 1700 1720
FT TRANSFAM 1720 1740
FT TRANSFAM 1740 1760
FT TRANSFAM 1760 1780
FT TRANSFAM 1780 1800
FT TRANSFAM 1800 1820
FT TRANSFAM 1820 1840
FT TRANSFAM 1840 1860
FT TRANSFAM 1860 1880
FT TRANSFAM 1880 1900
FT TRANSFAM 1900 1920
FT TRANSFAM 1920 1940
FT TRANSFAM 1940 1960
FT TRANSFAM 1960 1980
FT TRANSFAM 1980 2000
FT TRANSFAM 2000 2020
FT TRANSFAM 2020 2040
FT TRANSFAM 2040 2060
FT TRANSFAM 2060 2080
FT TRANSFAM 2080 2100
FT TRANSFAM 2100 2120
FT TRANSFAM 2120 2140
FT TRANSFAM 2140 2160
FT TRANSFAM 2160 2180
FT TRANSFAM 2180 2200
FT TRANSFAM 2200 2220
FT TRANSFAM 2220 2240
FT TRANSFAM 2240 2260
FT TRANSFAM 2260 2280
FT TRANSFAM 2280 2300
FT TRANSFAM 2300 2320
FT TRANSFAM 2320 2340
FT TRANSFAM 2340 2360
FT TRANSFAM 2360 2380
FT TRANSFAM 2380 2400
FT TRANSFAM 2400 2420
FT TRANSFAM 2420 2440
FT TRANSFAM 2440 2460
FT TRANSFAM 2460 2480
FT TRANSFAM 2480 2500
FT TRANSFAM 2500 2520
FT TRANSFAM 2520 2540
FT TRANSFAM 2540 2560
FT TRANSFAM 2560 2580
FT TRANSFAM 2580 2600
FT TRANSFAM 2600 2620
FT TRANSFAM 2620 2640
FT TRANSFAM 2640 2660
FT TRANSFAM 2660 2680
FT TRANSFAM 2680 2700
FT TRANSFAM 2700 2720
FT TRANSFAM 2720 2740
FT TRANSFAM 2740 2760
FT TRANSFAM 2760 2780
FT TRANSFAM 2780 2800
FT TRANSFAM 2800 2820
FT TRANSFAM 2820 2840
FT TRANSFAM 2840 2860
FT TRANSFAM 2860 2880
FT TRANSFAM 2880 2900
FT TRANSFAM 2900 2920
FT TRANSFAM 2920 2940
FT TRANSFAM 2940 2960
FT TRANSFAM 2960 2980
FT TRANSFAM 2980 3000
FT TRANSFAM 3000 3020
FT TRANSFAM 3020 3040
FT TRANSFAM 3040 3060
FT TRANSFAM 3060 3080
FT TRANSFAM 3080 3100
FT TRANSFAM 3100 3120
FT TRANSFAM 3120 3140
FT TRANSFAM 3140 3160
FT TRANSFAM 3160 3180
FT TRANSFAM 3180 3200
FT TRANSFAM 3200 3220
FT TRANSFAM 3220 3240
FT TRANSFAM 3240 3260
FT TRANSFAM 3260 3280
FT TRANSFAM 3280 3300
FT TRANSFAM 3300 3320
FT TRANSFAM 3320 3340
FT TRANSFAM 3340 3360
FT TRANSFAM 3360 3380
FT TRANSFAM 3380 3400
FT TRANSFAM 3400 3420
FT TRANSFAM 3420 3440
FT TRANSFAM 3440 3460
FT TRANSFAM 3460 3480
FT TRANSFAM 3480 3500
FT TRANSFAM 3500 3520
FT TRANSFAM 3520 3540
FT TRANSFAM 3540 3560
FT TRANSFAM 3560 3580
FT TRANSFAM 3580 3600
FT TRANSFAM 3600 3620
FT TRANSFAM 3620 3640
FT TRANSFAM 3640 3660
FT TRANSFAM 3660 3680
FT TRANSFAM 3680 3700
FT TRANSFAM 3700 3720
FT TRANSFAM 3720 3740
FT TRANSFAM 3740 3760
FT TRANSFAM 3760 3780
FT TRANSFAM 3780 3800
FT TRANSFAM 3800 3820
FT TRANSFAM 3820 3840
FT TRANSFAM 3840 3860
FT TRANSFAM 3860 3880
FT TRANSFAM 3880 3900
FT TRANSFAM 3900 3920
FT TRANSFAM 3920 3940
FT TRANSFAM 3940 3960
FT TRANSFAM 3960 3980
FT TRANSFAM 3980 4000
FT TRANSFAM 4000 4020
FT TRANSFAM 4020 4040
FT TRANSFAM 4040 4060
FT TRANSFAM 4060 4080
FT TRANSFAM 4080 4100
FT TRANSFAM 4100 4120
FT TRANSFAM 4120 4140
FT TRANSFAM 4140 4160
FT TRANSFAM 4160 4180
FT TRANSFAM 4180 4200
FT TRANSFAM 4200 4220
FT TRANSFAM 4220 4240
FT TRANSFAM 4240 4260
FT TRANSFAM 4260 4280
FT TRANSFAM 4280 4300
FT TRANSFAM 4300 4320
FT TRANSFAM 4320 4340
FT TRANSFAM 4340 4360
FT TRANSFAM 4360 4380
FT TRANSFAM 4380 4400
FT TRANSFAM 4400 4420
FT TRANSFAM 4420 4440
FT TRANSFAM 4440 4460
FT TRANSFAM 4460 4480
FT TRANSFAM 4480 4500
FT TRANSFAM 4500 4520
FT TRANSFAM 4520 4540
FT TRANSFAM 4540 4560
FT TRANSFAM 4560 4580
FT TRANSFAM 4580 4600
FT TRANSFAM 4600 4620
FT TRANSFAM 4620 4640
FT TRANSFAM 4640 4660
FT TRANSFAM 4660 4680
FT TRANSFAM 4680 4700
FT TRANSFAM 4700 4720
FT TRANSFAM 4720 4740
FT TRANSFAM 4740 4760
FT TRANSFAM 4760 4780
FT TRANSFAM 4780 4800
FT TRANSFAM 4800 4820
FT TRANSFAM 4820 4840
FT TRANSFAM 4840 4860
FT TRANSFAM 4860 4880
FT TRANSFAM 4880 4900
FT TRANSFAM 4900 4920
FT TRANSFAM 4920 4940
FT TRANSFAM 4940 4960
FT TRANSFAM 4960 4980
FT TRANSFAM 4980 5000
FT TRANSFAM 5000 5020
FT TRANSFAM 5020 5040
FT TRANSFAM 5040 5060
FT TRANSFAM 5060 5080
FT TRANSFAM 5080 5100
FT TRANSFAM 5100 5120
FT TRANSFAM 5120 5140
FT TRANSFAM 5140 5160
FT TRANSFAM 5160 5180
FT TRANSFAM 5180 5200
FT TRANSFAM 5200 5220
FT TRANSFAM 5220 5240
FT TRANSFAM 5240 5260
FT TRANSFAM 5260 5280
FT TRANSFAM 5280 5300
FT TRANSFAM 5300 5320
FT TRANSFAM 5320 5340
FT TRANSFAM 5340 5360
FT TRANSFAM 5360 5380
FT TRANSFAM 5380 5400
FT TRANSFAM 5400 5420
FT TRANSFAM 5420 5440
FT TRANSFAM 5440 5460
FT TRANSFAM 5460 5480
FT TRANSFAM 5480 5500
FT TRANSFAM 5500 5520
FT TRANSFAM 5520 5540
FT TRANSFAM 5540 5560
FT TRANSFAM 5560 5580
FT TRANSFAM 5580 5600
FT TRANSFAM 5600 5620
FT TRANSFAM 5620 5640
FT TRANSFAM 5640 5660
FT TRANSFAM 5660 5680
FT TRANSFAM 5680 5700
FT TRANSFAM 5700 5720
FT TRANSFAM 5720 5740
FT TRANSFAM 5740 5760
FT TRANSFAM 5760 5780
FT TRANSFAM 5780 5800
FT TRANSFAM 5800 5820
FT TRANSFAM 5820 5840
FT TRANSFAM 5840 5860
FT TRANSFAM 5860 5880
FT TRANSFAM 5880 5900
FT TRANSFAM 5900 5920
FT TRANSFAM 5920 5940
FT TRANSFAM 5940 5960
FT TRANSFAM 5960 5980
FT TRANSFAM 5980 6000
FT TRANSFAM 6000 6020
FT TRANSFAM 6020 6040
FT TRANSFAM 6040 6060
FT TRANSFAM 6060 6080
FT TRANSFAM 6080 6100
FT TRANSFAM 6100 6120
FT TRANSFAM 6120 6140
FT TRANSFAM 6140 6160
FT TRANSFAM 6160 6180
FT TRANSFAM 6180 6200
FT TRANSFAM 6200 6220
FT TRANSFAM 6220 6240
FT TRANSFAM 6240 6260
FT TRANSFAM 6260 6280
FT TRANSFAM 6280 6300
FT TRANSFAM 6300 6320
FT TRANSFAM 6320 6340
FT TRANSFAM 6340 6360
FT TRANSFAM 6360 6380
FT TRANSFAM 6380 6400
FT TRANSFAM 6400 6420
FT TRANSFAM 6420 6440
FT TRANSFAM 6440 6460
FT TRANSFAM 6460 6480
FT TRANSFAM 6480 6500
FT TRANSFAM 6500 6520
FT TRANSFAM 6520 6540
FT TRANSFAM 6540 6560
FT TRANSFAM 6560 6580
FT TRANSFAM 6580 6600
FT TRANSFAM 6600 6620
FT TRANSFAM 6620 6640
FT TRANSFAM 6640 6660
FT TRANSFAM 6660 6680
FT TRANSFAM 6680 6700
FT TRANSFAM 6700 6720
FT TRANSFAM 6720 6740
FT TRANSFAM 6740 6760
FT TRANSFAM 6760 6780
FT TRANSFAM 6780 6800
FT TRANSFAM 6800 6820
FT TRANSFAM 6820 6840
FT TRANSFAM 6840 6860
FT TRANSFAM 6860 6880
FT TRANSFAM 6880 6900
FT TRANSFAM 6900 6920
FT TRANSFAM 6920 6940
FT TRANSFAM 6940 6960
FT TRANSFAM 6960 6980
FT TRANSFAM 6980 7000
FT TRANSFAM 7000 7020
FT TRANSFAM 7020 7040
FT TRANSFAM 7040 7060
FT TRANSFAM 7060 7080
FT TRANSFAM 7080 7100
FT TRANSFAM 7100 7120
FT TRANSFAM 7120 7140
FT TRANSFAM 7140 7160
FT TRANSFAM 7160 7180
FT TRANSFAM 7180 7200
FT TRANSFAM 7200 7220
FT TRANSFAM 7220 7240
FT TRANSFAM 7240 7260
FT TRANSFAM 7260 7280
FT TRANSFAM 7280 7300
FT TRANSFAM 7300 7320
FT TRANSFAM 7320 7340
FT TRANSFAM 7340 7360
FT TRANSFAM 7360 7380
FT TRANSFAM 7380 7400
FT TRANSFAM 7400 7420
FT TRANSFAM 7420 7440
FT TRANSFAM 7440 7460
FT TRANSFAM 7460 7480
FT TRANSFAM 7480 7500
FT TRANSFAM 7500 7520
FT TRANSFAM 7520 7540
FT TRANSFAM 7540 7560
FT TRANSFAM 7560 7580
FT TRANSFAM 7580 7600
FT TRANSFAM 7600 7620
FT TRANSFAM 7620 7640
FT TRANSFAM 7640 7660
FT TRANSFAM 7660 7680
FT TRANSFAM 7680 7700
FT TRANSFAM 7700 7720
FT TRANSFAM 7720 7740
FT TRANSFAM 7740 7760
FT TRANSFAM 7760 7780
FT TRANSFAM 7780 7800
FT TRANSFAM 7800 7820
FT TRANSFAM 7820 7840
FT TRANSFAM 7840 7860
FT TRANSFAM 7860 7880
FT TRANSFAM 7880 7900
FT TRANSFAM 7900 7920
FT TRANSFAM 7920 7940
FT TRANSFAM 7940 7960
FT TRANSFAM 7960 7980
FT TRANSFAM 7980 8000
FT TRANSFAM 8000 8020
FT TRANSFAM 8020 8040
FT TRANSFAM 8040 8060
FT TRANSFAM 8060 8080
FT TRANSFAM 8080 8100
FT TRANSFAM 8100 8120
FT TRANSFAM 8120 8140
FT TRANSFAM 8140 8160
FT TRANSFAM 8160 8180
FT TRANSFAM 8180 8200
FT TRANSFAM 8200 8220
FT TRANSFAM 8220 8240
FT TRANSFAM 8240 8260
FT TRANSFAM 8260 8280
FT TRANSFAM 8280 8300
FT TRANSFAM 8300 8320
FT TRANSFAM 8320 8340
FT TRANSFAM 8340 8360
FT TRANSFAM 8360 8380
FT TRANSFAM 8380 8400
FT TRANSFAM 8400 8420
FT TRANSFAM 8420 8440
FT TRANSFAM 8440 8460
FT TRANSFAM 8460 8480
FT TRANSFAM 8480 8500
FT TRANSFAM 8500 8520
FT TRANSFAM 8520 8540
FT TRANSFAM 8540 8560
FT TRANSFAM 8560 8580
FT TRANSFAM 8580 8600
FT TRANSFAM 8600 8620
FT TRANSFAM 8620 8640
FT TRANSFAM 8640 8660
FT TRANSFAM 8660 8680
FT TRANSFAM 8680 8700
FT TRANSFAM 8700 8720
FT TRANSFAM 8720 8740
FT TRANSFAM 8740 8760
FT TRANSFAM 8760 8780
FT TRANSFAM 8780 8800
FT TRANSFAM 8800 8820
FT TRANSFAM 8820 8840
FT TRANSFAM 8840 8860
FT TRANSFAM 8860 8880
FT TRANSFAM 8880 8900
FT TRANSFAM 8900 8920
FT TRANSFAM 8920 8940
FT TRANSFAM 8940 8960
FT TRANSFAM 8960 8980
FT TRANSFAM 8980 9000
FT TRANSFAM 9000 9020
FT TRANSFAM 9020 9040
FT TRANSFAM 9040 9060
FT TRANSFAM 9060 9080
FT TRANSFAM 9080 9100
FT TRANSFAM 9100 9120
FT TRANSFAM 9120 9140
FT TRANSFAM 9140 9160
FT TRANSFAM 9160 9180
FT TRANSFAM 9180 9200
FT TRANSFAM 9200 9220
FT TRANSFAM 9220 9240
FT TRANSFAM 9240 9260
FT TRANSFAM 9260 9280
FT TRANSFAM 9280 9300
FT TRANSFAM 9300 9320
FT TRANSFAM 9320 9340
FT TRANSFAM 9340 9360
FT TRANSFAM 9360 9380
FT TRANSFAM 9380 9400
FT TRANSFAM 9400 9420
FT TRANSFAM 9420 9440
FT TRANSFAM 9440 9460
FT TRANSFAM 9460 9480
FT TRANSFAM 9480 9500
FT TRANSFAM 9500 9520
FT TRANSFAM 9520 9540
FT TRANSFAM 9540 9560
FT TRANSFAM 9560 9580
FT TRANSFAM 9580 9600
FT TRANSFAM 9600 9620
FT TRANSFAM 9620 9640
FT TRANSFAM 9640 9660
FT TRANSFAM 9660 9680
FT TRANSFAM 9680 9700
FT TRANSFAM 9700 9720
FT TRANSFAM 9720 9740
FT TRANSFAM 9740 9760
FT TRANSFAM 9760 9780
FT TRANSFAM 9780 9800
FT TRANSFAM 9800 9820
FT TRANSFAM 9820 9840
FT TRANSFAM 9840 9860
FT TRANSFAM 9860 9880
FT TRANSFAM 9880 9900
FT TRANSFAM 9900 9920
FT TRANSFAM 9920 9940
FT TRANSFAM 9940 9960
FT TRANSFAM 9960 9980
FT TRANSFAM 9980 10000

```

```

DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Zinc finger protein GLI2 (Tax helper protein).
GN GLI2 OR THP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX MEDLINE=98216760; PubMed=9557682;
RA Tanimura A., Dan S., Yoshida M.;
RT "Cloning of novel isoforms of the human GLI2 oncogene and their
RT activities to enhance tax-dependent transcription of the human T-cell
RT leukemia virus type 1 genome."
RT J. Virol. 72:3958-3964(1998).
RN [2]
RP SEQUENCE OF 1-529 FROM N.A.
RX MEDLINE=9353614; PubMed=8350401;
RA Tanimura A., Teshima H., Fujisawa J., Yoshida M.;
RT "A new regulatory element that augments the Tax-dependent enhancer of
RT human T-cell leukemia virus type 1 and cloning of cDNAs encoding its
RT binding proteins."
RT J. Virol. 67:5375-5382(1993).
RN [3]
RP SEQUENCE OF 129-233 FROM N.A.
RX MEDLINE=89096896; PubMed=2850480;
RA Ruppert J.M., Kinzler K.W., Wong A.J., Bigner S.H., Kao F.T.,
RA Law M.L., Seanez H.N., O'Brien S.J., Vogelstein B.;
RT "The Gli-Kruppel family of human genes."
RT Mol. Cell. Biol. 8:3104-3113(1988).
CC -1- FUNCTION: MAY PLAY A ROLE DURING EMBRYOGENESIS. BINDS TO THE DNA
CC SEQUENCE 5'-GACCACCA-3' WHICH IS PART OF THE TRE-2S REGULATORY
CC ELEMENT THAT AUGMENTS THE TAX-DEPENDENT ENHANCER OF HUMAN T-CELL
CC LEUKEMIA VIRUS TYPE 1. IMPLICATED IN THE TRANSDUCTION OF SHH
CC SIGNAL.
CC -1- SUBCELLULAR LOCATION: Nuclear (potential).
CC -1- ALTERNATIVE PRODUCTS: 4 ISOFORMS: ALPHA (SHOWN HERE), BETA, GAMMA
CC AND DELTA; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- SIMILARITY: BELONGS TO THE GLI FAMILY OF C2H2-TYPE ZINC-FINGER
CC PROTEINS.
CC -1- CAUTION: REF. 2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO FRAMESHIFTS
CC IN POSITION 427 AND 433.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AB007295; BAA25665.1; -
DR EMBL: AB007296; BAA25666.1; -
DR EMBL: AB007297; BAA25667.1; -
DR EMBL: AB007298; BAA25668.1; -
DR EMBL: AB007299; BAA25669.1; -
DR EMBL: D14828; BAA03569.1; ALT_FRAME.
DR EMBL: D14827; BAA03568.1; ALT_FRAME.
DR EMBL: M20672; -; NOT_ANNOTATED_CDS.
DR EMBL: M20673; AAA35898.1; -
DR PIR: A31201; A31201.
DR HSP: P08151; 2GL1.
DR MIM: 165230; -
DR InterPro: IPR000822; Znf-C2H2.
DR Pfam: PF00096; Zf-C2H2_5.
DR PRINTS: PRO0048; ZINC_FINGER.
DR SMART: SM00355; Znf-C2H2; 5.
DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 4.
DR PROSITE: PS00157; ZINC_FINGER_C2H2_2; 4.
KW Transcription regulation; Zinc-finger; Metal-binding; DNA-binding;
KW Nuclear protein; Repeat; Alternative splicing.
FT DOMAIN 109 261
FT ZN_FING 109 134
FT ZN_FING 142 169
FT C2H2-TYPE.

```

```

FT 2N_FING 175 199 C2H2-TYPE.
FT 2N_FING 205 230 C2H2-TYPE.
FT 2N_FING 236 261 C2H2-TYPE.
FT VARSPPLIC 67 83 MISSING (IN ISOFORM BETA AND ISOFORM
FT VARSPPLIC 821 829 MISSING (IN ISOFORM GAMMA
FT VARSPPLIC 830 1258 MISSING (IN ISOFORM GAMMA AND ISOFORM
FT VARSPPLIC 830 1258 MISSING (IN ISOFORM GAMMA AND ISOFORM
FT CONFLICT 390 391 OL -> HV (IN REF. 2).
FT SEQUENCE 1258 AA, 133309 MW, 107CE477D36603C CRC64;

Query Match
Best Local Similarity 27.7%; Score 60.5; DB 1; Length 1258;
Matches 18; Conservative 9; Mismatches 19; Indels 19; Gaps 3;

OY 4 HKLAHRLADPLDLSLRPGH-----APHPRRDGGGIDSCV-----QSDG 46
DB 1073 HAMAMPSSQE--TAAEAVPKGAMGMSVPPQPPQDAGAPDHSLYYGQIHMYEDQG 1130
OY 47 VLQHL 51
DB 1131 GLENL 1135

RESULT 12
BAT2_HUMAN
ID BAT2_HUMAN STANDARD: PRT; 2142 AA.
AC P48634;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Large proline-rich protein Bat2 (HLA-B-associated transcript 2).
GN Bat2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-T-cell;
RX MEDLINE=90192810; PubMed=2156268;
RA Banerji J., Sands J., Strominger J.L., Spies T.;
RT "A gene pair from the human major histocompatibility complex encodes
RT large proline-rich proteins with multiple repeated motifs and a
RT single ubiquitin-like domain.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:2374-2378(1990).
RN [2]
RP SEQUENCE OF 1-1860 FROM N.A.
RX MEDLINE=93272029; PubMed=8499947;
RA Irls F.J.M., Bouguetieret L., Prieur S., Caterina D., Primas G.,
RA Perrot V., Jurka J., Rodriguez-Tome P., Claverie J.-M., Dausset J.,
RA Cohen D.;
RT "Dense ALU clustering and a potential new member of the NF kappa B
RT family within a 90 kilobase HLA class III segment.";
RL Nat. Genet. 3:137-145(1993).
CC -1- FUNCTION: UNKNOWN.
CC -1- TISSUE SPECIFICITY: LIMITED TO CELL-LINES OF LEUKEMIC ORIGIN.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M33509; AAA35585.1; -
CC EMBL: M33518; AAA35586.1; -
CC EMBL: M33512; AAA35586.1; JOINED.
CC EMBL: Z15025; CAA78744.1; -
CC PIR: B35098; B35098.
CC PIR: S36152; S36152.

```

```

DR Repeat.
KV Repeat.
FT DOMAIN 519 524 POLY-PRO.
FT DOMAIN 636 657 GLN-RICH.
FT DOMAIN 684 688 POLY-PRO.
FT DOMAIN 699 704 POLY-PRO.
FT DOMAIN 814 821 POLY-PRO.
FT DOMAIN 1340 1345 POLY-GLY.
FT DOMAIN 1398 1403 POLY-GLY.
FT DOMAIN 1436 1442 POLY-PRO.
FT DOMAIN 1482 1491 POLY-PRO.
FT DOMAIN 1795 1795 4 X 57 AA TYPE A REPEATS.
FT REPEAT 41 95 1-1.
FT REPEAT 98 154 1-2.
FT REPEAT 281 337 1-3.
FT REPEAT 1740 1795 1-4.
FT REPEAT 337 549 2 X TYPE B REPEATS.
FT REPEAT 418 418 2-1.
FT REPEAT 476 549 2-2.
FT DOMAIN 1899 2089 3 X 50 AA TYPE C REPEATS.
FT REPEAT 1899 1948 3-1.
FT REPEAT 1965 2014 3-2.
FT REPEAT 2040 2089 3-3.
FT CONFLICT 57 57 R -> A (IN REF. 2).
FT CONFLICT 109 109 Q -> S (IN REF. 2).
FT CONFLICT 414 414 P -> PPHRGAGNMGPP (IN REF. 2).
FT CONFLICT 532 532 T -> K (IN REF. 2).
FT CONFLICT 682 682 Q -> K (IN REF. 2).
FT CONFLICT 730 730 E -> D (IN REF. 2).
FT CONFLICT 750 750 L -> R (IN REF. 2).
FT CONFLICT 834 834 G -> T (IN REF. 2).
FT CONFLICT 1035 1035 M -> L (IN REF. 2).
FT CONFLICT 1068 1068 P -> R (IN REF. 2).
FT CONFLICT 1285 1285 G -> A (IN REF. 2).
FT CONFLICT 1400 1400 G -> A (IN REF. 2).
FT CONFLICT 1611 1611 T -> S (IN REF. 2).
FT CONFLICT 1729 1729 G -> A (IN REF. 2).
SQ SEQUENCE 2142 AA, 227840 MW, 32DFF1B9B52420A CRC64;

Query Match
Best Local Similarity 17.9%; Score 60.5; DB 1; Length 2142;
Matches 14; Conservative 7; Mismatches 15; Indels 3; Gaps 1;

OY 22 PRGHAPHPRRDGGGSDGVSDVQLHQRPHVKG 60
DB 1492 POGSPRPRTFYRQVNSGLSSD--PHFEPPGVMRG 1527

RESULT 13
TOP1_MYCTU
ID TOP1_MYCTU STANDARD: PRT; 934 AA.
AC Q59567; Q08383;
DT 01-NOV-1997 (Rel. 35, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA topoisomerase I (EC 5.99.1.2) (Omega-protein) (Relaxing enzyme)
DE (unwinding enzyme) (Swivelase).
DE TOPO OR RV3646C OR MT3749 OR MTCY15C10.06.
GN Mycobacterium tuberculosis.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garrier T., Churcher C., Harris D.,
RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Fellwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jajels K., Kirogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,

```

RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence."
 RL Nature 393:537-544(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CDC 1551 / Oshkosh;
 RA Fleischman R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
 RA Delcher A., Uterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains."
 RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 1-895 FROM N.A.
 RC STRAIN-ERDMANN;
 RX MEDLINE=97080548; PubMed=8921893;
 RA Yang F.D., Lu G., Rubin H.;
 RT "Cloning, expression, purification and characterization of DNA
 RT topoisomerase I of Mycobacterium tuberculosis."
 RL Gene 178:63-69(1996).
 CC -1- FUNCTION: THE REACTION CATALYZED BY TOPOISOMERASES LEADS TO THE
 CC CONVERSION OF ONE TOPOLOGICAL ISOMER OF DNA TO ANOTHER.
 CC -1- CATALYTIC ACTIVITY: ATP-independent breakage of single-stranded
 CC DNA, followed by passage and rejoining.
 CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
 CC -1- MISCELLANEOUS: WHEN A TOPOISOMERASE TRANSIENTLY BREAKS A DNA
 CC BACKBONE BOND, IT SIMULTANEOUSLY FORMS A PROTEIN-DNA LINK, IN
 CC WHICH A TYROSYL OXYGEN IN THE ENZYME IS JOINED TO A DNA PHOSPHORUS
 CC AT ONE END OF THE ENZYME-SEVERED DNA STRAND.
 CC -1- SIMILARITY: BELONGS TO PROKARYOTIC TYPE I/III TOPOISOMERASE
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Z95436; CAB08843.1; -;
 DR EMBL: AE007173; AAK48109.1; -;
 DR EMBL: U40159; AAC4599.1; -;
 DR HSSP: P06612; 1ECL.
 DR TIGR: MT3749; -;
 DR TubercuList: Rv3646c; -;
 DR InterPro: IPR003601; DNATopI_ATP_bind.
 DR InterPro: IPR003602; DNATopI_DNA_bind.
 DR InterPro: IPR000380; Pro_topoisomerase.
 DR InterPro: IPR002936; Toprim.
 DR Pfam: PF01131; Topoisom_bac; 1.
 DR Pfam: PF01751; Toprim; 1.
 DR PRINTS: PR00417; PRTPIKRASEI.
 DR SMART: SM00437; TOP1AC; 1.
 DR SMART: SM00436; TOP1BC; 1.
 DR SMART: SM00493; TOPRIM; 1.
 DR PROSITE: PS00396; TOPOISOMERASE_1_PROK; 1.
 KW Isomerase; Topoisomerase; DNA-binding; Complete proteome.
 FT ACT_SITE 342 342 DNA CLEAVAGE (BY SIMILARITY).
 FT CONFLICT 1 1 M -> MERGAOL (IN REF. 3).
 FT CONFLICT 53 61 PRAASDVPA -> RGPACTR (IN REF. 3).
 FT CONFLICT 255 281 RDFSIGLTRKGDDEVIVDESGATLA -> AISTRMARCA
 FT KATKISCTRGARPPWP (IN REF. 3).
 FT EYVA -> GYVR (IN REF. 3).
 FT G -> V (IN REF. 3).
 FT PURLGIDDPASGKPPVINDGFGFPVYTDGETNASTLRKGDV
 FT AS -> ACASWEQIRRASQSSRTADSGRTSPVTRPMPAC
 FT VRATTWLP (IN REF. 3).
 FT SEQUENCE 934 AA: 102335 MW: B29E2C17897781BA CRC64;

Query Match 17.5%; Score 59; DB 1; Length 934;
 Best Local Similarity 30.6%; Pred. No. 49;
 Matches 22; Conservative 11; Mismatches 25; Indels 14; Gaps 3;
 QY 3 DKKLAHRPALQDPLQSLRPRGHAPPPR-----DIAGGLDSGVSDGYLQHL 51
 DB 487 ERLPLRPGQRLDIVELPDPGHATNPARYTEASLWKALELIGIRPSTYSS--IIKTI 544
 QY 52 QRGHW-KIGTA 62
 DB 545 QDRGIVHKKGSA 556
 RESULT 14
 ID SALLM_DROME STANDARD; PRT; 1355 AA.
 AC P39770;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Homeotic protein spalt-major.
 GN SALLM.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxId=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94139659; PubMed=7905822;
 RA Kuehnlein R.P., Frommer G., Friedrich M., Gonzalez-Gaitan M.,
 RA Weber A., Wagner-Bernholz J.F., Gehring W.J., Jaekle H., Schuh R.;
 RT "spalt encodes an evolutionarily conserved zinc finger protein of
 RT novel structure which provides homeotic gene function in the head and
 RT tail region of the Drosophila embryo."
 RL EMBO J. 13:168-179(1994).
 CC -1- FUNCTION: REQUITED FOR THE ESTABLISHMENT OF THE POSTERIOR-MOST
 CC HEAD AND THE ANTERIOR-MOST TAIL SEGMENTS OF THE EMBRYO. PROBABLY
 CC FUNCTION AS A TRANSCRIPTIONAL REGULATOR. COULD REPRESS THE
 CC TRANSCRIPTION OF THE TSH GENE.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- DEVELOPMENTAL STAGE: FIRST EXPRESSED AT BLASTODERM STAGE AND
 CC LATER IN RESTRICTED AREAS OF THE EMBRYONIC NERVOUS SYSTEM AS
 CC WELL AS IN THE DEVELOPING TRACHEA.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X75541; CAA53229.1; -;
 DR HSSP: P15822; 1BBO.
 DR FlyBase: FBgn004579; salm.
 DR InterPro: IPR000822; Znf-C2H2.
 DR Pfam: PF00096; ZF-C2H2; 7.
 DR PRINTS: PR00048; ZINCINGER.
 DR SMART: SM00355; ZNF_C2H2; 7.
 DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 7.
 DR PROSITE: PS50157; ZINC_FINGER_C2H2_2; 7.
 KW Developmental protein; Zinc-finger; Metal-binding; DNA-binding;
 KW Nuclear protein; Transcription regulation; Repeat.
 FT ZN_FING 451 473 C2H2-TYPE.
 FT ZN_FING 479 501 C2H2-TYPE.
 FT ZN_FING 824 846 C2H2-TYPE.
 FT ZN_FING 852 874 C2H2-TYPE.
 FT ZN_FING 884 906 C2H2-TYPE.
 FT ZN_FING 1289 1311 C2H2-TYPE.
 FT ZN_FING 1317 1339 C2H2-TYPE.
 FT SEQUENCE 1355 AA: 148995 MW: 6FC4EBDC0BC6355E CRC64;

THIS PAGE BLANK (USPTO)

DT	01-MAY-2000 (TREMBlrel. 13, last sequence update)
DT	01-OCT-2001 (TREMBlrel. 18, last annotation update)
DE	CG4124 PROTEIN.
GN	CG4124.
OS	Drosophila melanogaster (Fruit fly).
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta.
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC	Ephyridioidea; Drosophilidae; Drosophila.
OX	NCBI_TaxID=7227;
RN	(1)
RP	SEQUENCE FROM N.A.
RC	STRAIN-BERKELEY;
RX	MEDLINE=20196006; PubMed=10731132;
RA	Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA	Brandon R.C., Rogers Y.H.C., Blaise R.G., Champe M., Pfeiffer B.D.,
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA	Abri J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA	Bailey R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA	Beeson K.Y., Benos P.V., Bertman B.P., Bhandari D., Bolshakov S.,
RA	Bortova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA	Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA	Durkin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA	Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA	Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA	Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA	Hoslin D., Houston K.A., Howland T.J., Wei M.-H., Ibeagwam C.,
RA	Jalali B.E., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA	Laslo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA	Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA	Mervulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA	Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA	Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pauley J.M.,
RA	Palazzo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA	Rehert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA	Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA	Spieler E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA	Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA	Wang Z.-Y., Weissarman D.A., Weinstein G.M., Weisenbach J.,
RA	Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA	Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA	Zheng X.H., Zhong F.N., Zhong Y., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA	Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT	"The genome sequence of Drosophila melanogaster.";
RL	Science 287:2185-2195(2000).
DR	EMBL: AE003588: AAF51439.1.
DR	EMBL: FB00031291: CG4124.
DR	InterPro: IPR000571: Zf-CCCH.
DR	Pfam: PF00642: Zf-CCCH; 1.
DR	SMART: SM00356; Znf_C3H1; 1.
SQ	SEQUENCE 1071 AA; 113488 MW; 0E1FE584E1A95132 CRC64;
Query Match	22.5%; Score 76; DB 5; Length 1071;
Best Local Similarity	35.6%; Pred. No. 1.5;
Matches 16; Conservative 7; Mismatches 22; Indels 0; Gaps 0;	
OY	15 PIIQSLRPGHAPHPRLDGGGGLGVGSDVDLHLORPHVKL 59
DB	456 PALMAVPSLKRPPRPDLSTDKAKAKIDINNVAGHVERPAAKL 500
RESULT 3	
ID	O9D5M3 PRELIMINARY; PRT: 234 AA.
AC	O9D5M3;
DT	01-JUN-2001 (TREMBlrel. 17, created)
DT	01-JUN-2001 (TREMBlrel. 17, last sequence update)

[illegible]

DR EMBL: AL136559; CAB66654.1; -.
DR HSPD: Q99418; IPBV.
DR InterPro: IPR001845; PH.
DR InterPro: IPR000904; Sec7.
DR InterPro: IPR01605; Spectrin_PH.
DR Pfam: PF00169; PH: 1.
DR PRINTS: PR00683; SPECTRINPH.
DR SMART: SM00233; PH: 1.
DR SMART: SM00222; Sec7: 1.
DR PROSITE: PSS0003; PH DOMAIN: 1.
DR PROSITE: PSS0190; SEC7: 1.
KW Hypothetical protein.
SQ SEQUENCE 771 AA; 84660 MW; 3D7E08229429ED78 CRC64;

Query Match	20.1%	Score 68	DB 4	Length 771
Best Local Similarity	34.6%	Pred. NO. 9.6		
Matches 18	Conservative 10	Mismatches 20	Indels 4	Gaps 2

0y 13 DDPLSLGSRPRGA---P HRRRLDGGGLSDGSQSDVGLNHLRPRGNVLGT 61
 : | | :
Db 3 EDKLLSAVPEDGDTRDPGEPEE-EPGVNRNGMASEGGLSSLCSGPHNERGT 53

RESULT 5
Q9Y4B5
ID Q9Y4B5 PRELIMINARY; PRT; 1353 AA.

DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE KIAA0802 PROTEIN (FRAGMENT).
GN KIAA0802.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid:9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
PC MEDLINE=99087487; PubMed=9872452;
RA Nagase T., Ishikawa K., Suyama M., Kikuno R., Miyajima N., Tanaka A.,
RA Kotani H., Nomura N., Ohara O.,
RT "Prediction of the coding sequences of unidentified human genes. XI.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 5:277-286(1998).
DR EMBL; AB018345; BAA34552.1; -
DR InterPro; IPR000100; Ribonuclease_P.
FT NON_TER 1 1
SQ SEQUENCE 1353 AA; 150386 MW; 27A38F75B6EBD03F CRC64;

Query Match	20.1%	Score 68	DB 4	Length 1353
Best Local Similarity	30.9%	Pred. No. 17		
Matches 17; Conservative	10;	Mismatches	22;	Indels 6;
				Gaps 2;

Qy 8 HRPALQDPLQSLRPRGHAPHPPPR-DLGGGLDSGVQSDGVQLHLQRPCHVKLGT 61
| : ||| : || : ||| : ||| : | : : :
Db 1219 HSPVQDPQKGLRAGSRSRSAEPPELGPQGETNSRG-----RSPDPGVGS 1268

RESULT 6	
09JMC1	
ID	09JMC1
AC	09JMC1
DT	01-OCT-2000 (TREMBlrel. 15, Created)
DT	01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT	01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE	PROLINE-RICH INOSITOL PHOSPHOSPHATE 5-PHOSPHATASE.
GN	PIPP.
OS	Rattus norvegicus (Rat).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus
XN	NCBI_TaxID=10116;

RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BRAIN;
 RX MEDLINE=20062911; PubMed=10593988;
 RA Mochizuki Y., Takenawa T.;
 RT "Novel hosttol polyporphosphate 5-phosphatase localizes at membrane
 ruffles.";
 RL J. Biol. Chem. 274:36790-36795(1999).
 DR EMBL; AB032551; BAA90553.1; -
 DR InterPro; IPR000300; IPPC.
 DR InterPro; IPR002965; P_f1ch_extensn.
 DR Pfam; PF00783; IPPC; 1-
 DR PRINTS; PRO1217; PRICHEXTENSN.
 DR SMART; SM00128; IPPC; 1.
 SO SEQUENCE 1001 AA; 107207 MW; 7BE7741FEE8F3FAB CRC64;

Query Match	20.0%;	Score	67.5;	DB	11;	Length	1001;
Best Local Similarity	38.6%;	Pred. No.	14;				
Matches	22;	Conservative	5;	Mismatches	15;	Indels	15;
						Gaps	4;

```

QY      14 DPLQ-SLRGRHAPH-----PPPR----DLGGI-DSGVSDGVLTQHTQRP 55
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      275 DERLSPSFARPEARHSPEDVLP RPPTQLPLDVSSSLPESGTSFGLSPFRPG 331

```

RESULT 7
Q9NH70
ID Q9NH70 PRELIMINARY; PRT; 517 AA.

DT 01-OCT-2000 (TREMBLrel, 15, created)
 DT 01-OCT-2000 (TREMBLrel, 15, last sequence update)
 DT 01-OCT-2000 (TREMBLrel, 15, last annotation update)
 DE LIN-14 PROTEIN (FRAGMENT).
 OS Caenorhabditis vulgatis.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_TaxID=31233;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hong Y., Lee R.C., Ambros V.R.;
 RT "Structure and Function Analysis of LIN-14, a Temporal Regulator of
 RT Postembryonic Developmental Events in *Caenorhabditis elegans*.";
 RL Mol. Cell. Biol. 0:0-0(2000).
 FR EMBL: AF231036; AAF34229.1; -.
 FT NON_TER 517 517
 SQ SEQUENCE 517 AA: 57074 MW: A3F4AD754775BE4 CRC64:

Query Match	19.8%	Score 67	DB 5	Length 517
Best Local Similarity	34.5%	Pred. No. 8.3		
Matches 20; Conservative	7	Mismatches 25	Indels 6	Gaps 2

```

Oy      2 GDHKLAIHPALODELLQSLRPRGHAPRP---PRDLGGGIDSGVSDGVLQHLQRPG 55
          |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db      75 GQTTLIH--VMQSSLRNKRPIAHPTVYQLECGQLRTVLSSPIDNAIIMNHLMPG 130

```

RESULT	8
09UG71	
ID	09UG71
AC	09UG71;
DT	01-OCT-2000 (TREMBLrel. 15, Created)
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE	ORF2.
OS	TT virus.
OC	Viruses; ssDNA viruses; unclassified ssDNA viruses.
OX	NCHI_Taxid=68887;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=TTVTON-KC186;
RX	MEDLINE=20428649; PubMed=10971131;
RA	Yakahashi K., Hijikata M., Samokhvalov E.I., Mishiro S.;
RT	"Full or near full length nucleotide sequences of TT virus variants

RT (types SANBAN and YONBAN) and the TT virus-like mini virus."
 RL InterVirology 43:119-123(2000).
 DR EMBL: AB038623; BAA03588.1; -
 DR InterPro: IPR004118; TT_ORF2.
 DR Pfam: PF02957; TT_ORF2; 1.
 SQ SEQUENCE 128 AA; 13049 MW; B8B480D2A1198B0F CRC64;

Query Match 19.5%; Score 66; DB 12; Length 128;
 Best Local Similarity 34.9%; Pred. No. 2.5;
 Matches 22; Conservative 0; Mismatches 21; Indels 20; Gaps 3;

OY 10 PALDPLQLSLRRGHA-----PHEPRDLGGIDSGVQSDGLQLHQ 52
 DB 50 PRPPGGLDQNPESGPGGPPALPALPAPDEPPRR-RGGAGDGV--DGLAIAN 106
 OY 53 RPG 55
 DB 107 APG 109

RESULT 9

O9FWH4 PRELIMINARY: PRT; 460 AA.

AC O9FWH4 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE HYPOTHEICAL.50.2 KDA PROTEIN.
 GN OSJNBAA0055P24.5.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzeae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. NIPPONBARE;
 RA Buehl C.R., Yuan Q., Moffat K.S., Hill J.N., Jenkins C.N., Burr P.C.,
 RA Hsiao J., Zismann V., Pai G., Bowman C.L., Fujii C.Y., Vanaken S.E.,
 RA Bowman C.L., Craven B., Uterback T.R., Khatai H., Feldblyum T.V.,
 RA Quackenbush J., White O., Salzberg S.L., Fraser C.M.;
 RT "Oryza sativa chromosome 10 BAC OSJNBAA0055P24 genomic sequence."
 RL Submitted (OCT-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AC037425; AAC13576.1; -
 DR InterPro: IPR002965; P_rich_extensn.
 DR PRINTS: PR01217; PRICEXTENSN.
 KW Hypothetical protein.
 SQ SEQUENCE 460 AA; 50230 MW; A9E30128E7CF97E1 CRC64;

Query Match 19.4%; Score 65.5; DB 10; Length 460;
 Best Local Similarity 53.6%; Pred. No. 11;
 Matches 15; Conservative 2; Mismatches 10; Indels 1; Gaps 1;

OY 8 HRPALQDPLQLSLRRGHA-PHPPRDL 34
 DB 72 HDPALRSLSLRSLPPAPAPHPHPPRHL 99

RESULT 10

O41979 PRELIMINARY: PRT; 230 AA.

AC O41979 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE HYPOTHEICAL.23.2 KDA PROTEIN.
 GN GAMMAHV.M12.
 OS Murid herpesvirus 4.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Gammaherpesvirinae.
 OX NCBI_TaxID=33708;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=WUMS;
 RX MEDLINE=97366649; PubMed=9223479;
 RA Virgin H.W. IV, Latreille P., Wamsley P., Hallsworth K., Weck K.E.,
 RA Dal Canto A.J., Speck S.H.;
 RT "Complete sequence and genomic analysis of murine gammaherpesvirus
 RT 68.";
 RL J. Virol. 71:5894-5904(1997).
 RN [2]
 RP SEQUENCE FROM N.A.

RC STRAIN=WUMS;
 RA Latreille P., Wamsley P., Waterston R.H.;
 RL Submitted (APR-1997) to the EMBL/Genbank/DBJ databases.
 DR EMBL: U97553; AAB66425.1; -
 KW Hypothetical protein.
 SQ SEQUENCE 230 AA; 23235 MW; 8091650176A6B880 CRC64;

Query Match 19.2%; Score 65; DB 12; Length 230;
 Best Local Similarity 38.0%; Pred. No. 6.1;
 Matches 19; Conservative 5; Mismatches 16; Indels 10; Gaps 2;

OY 6 LAHRPALDPP-----LQSLRRGHAHPHPPRDLGGIDSGVQSDGV 47
 DB 74 LSPRLALSPPPKKRGGGLISSLPARGSPYPP--LGRQAPTAAGGSGV 121

RESULT 11

O9BOA8 PRELIMINARY: PRT; 476 AA.

AC O9BOA8 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
 DE SIMILAR TO T54 PROTEIN.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LUNG CARCINOMA;
 RA Strausberg R.;
 RT Submitted (FEB-2001) to the EMBL/Genbank/DBJ databases.
 RL [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LUNG CARCINOMA;
 RA Strausberg R.;
 RL Submitted (NOV-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL: BC003148; AAH03148.1; -
 DR EMBL: BC000397; AAH00397.1; -
 DR InterPro: IPR000467; G_patch.
 DR Pfam: PF01585; G_patch; 1.
 DR SMART: SM00443; G_patch; 1.
 SQ SEQUENCE 476 AA; 52228 MW; 36P1C5771969897D CRC64;

Query Match 19.1%; Score 64.5; DB 4; Length 476;
 Best Local Similarity 33.3%; Pred. No. 15;
 Matches 15; Conservative 6; Mismatches 11; Indels 13; Gaps 1;

OY 17 LOSLRP-----GHAPPPRDLGGIDSGVQSDGV 48
 DB 57 LOSVAPQEQAPKELVLPILQNGHROPAPRPPGSDTGTALADGV 101

RESULT 12

O9VW86 PRELIMINARY: PRT; 603 AA.

AC O9VW86 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE CG13811 PROTEIN.
 GN CG13811.
 OS Drosophila melanogaster (Fruit fly).

SEQ SEQUENCE 785 AA; 87658 MW; 3C47EEF12B2F825B CRC64;

Job time : 29 secs

Query Match 19.1%; Score 64.5; DB 4; Length 785;

Best Local Similarity 28.6%; Pred. No. 25;

Matches 18; Conservative 10; Mismatches 24; Indels 11; Gaps 3;

OY 4 HKLAHRLADPLDLSLRPGHAPHP---PPRDLGGDLSGVSDGVLOHLQRPCHVKL 59
 | | | | | : | | | | | : | | | | | : | | | | |
 DB 632 HYLSHVPGILPNC-----QNTVYPPTFNLPPHSAVGSNDNDIPDLAIKH-SRPGPTAN 684

OY 60 GTA 62

DB 685 GAS 687

RESULT 15

ID Q9UHH6 PRELIMINARY; PRT; 1281 AA.
 AC Q9UHH6;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, last annotation update)
 DE TRPS1 (ZINC FINGER PROTEIN GC79).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Momeni P., Gloeckner G., Schmidt O.S., von Holtum D., Hennekam R.,
 RA Meunke P., Zabel B., Rosenthal A., Horsthemke B., Luedecke H.-J.;
 RT "Mutations in a novel zinc finger transcription factor gene on human
 RT chromosome 8q cause tricho-rhino-phalangeal syndrome type I (TRPS1).";
 RL Nat. Genet. 0:0-0(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20432376; PubMed=10974077;
 RA Chang G.T.G., Steenbeek M., Schippers E., Blok L.J., van Weerden W.M.,
 RA van Alewijk D.C., Eussen B.H., van Steenbrugge G.J., Brinkmann A.O.;
 RT "Characterization of a zinc-finger protein and its Association with
 RT Apoptosis in Prostate Cancer Cells.";
 RL J. Natl. Cancer Inst. 92:1414-1421(2000).
 DR EMBL: AF178030; AAF16000.1; -;
 DR EMBL: AF264784; AAG21134.1; -;
 DR HSSP: P17679; IGMF.
 DR InterPro: IPR000822; Znf-C2H2.
 DR InterPro: IPR000679; Znf-GATA.
 DR Pfam: PF00320; GATA.1.
 DR Pfam: PF00096; ZF-C2H2; 7.
 DR PRINTS: PR00619; GATAZNFINGER.
 DR SMART: SM00355; ZNF_C2H2; 7.
 DR SMART: SM00401; ZNF_GATA; 1.
 DR PROSITE: PS00344; GATA_ZN_FINGER_1; 1.
 DR PROSITE: PS0114; GATA_ZN_FINGER_2; 1.
 DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_2.
 DR PROSITE: PS0157; ZINC_FINGER_C2H2_2; 1.
 KW DNA-binding; zinc-finger.
 SQ SEQUENCE 1281 AA; 141520 MW; 2157B04F5BB71CC CRC64;

Query Match 19.1%; Score 64.5; DB 4; Length 1281;

Best Local Similarity 28.6%; Pred. No. 42;

Matches 18; Conservative 10; Mismatches 24; Indels 11; Gaps 3;

OY 4 HKLAHRLADPLDLSLRPGHAPHP---PPRDLGGDLSGVSDGVLOHLQRPCHVKL 59
 | | | | | : | | | | | : | | | | | : | | | | |
 DB 1126 HYLSHVPGILPNC-----QNTVYPPTFNLPPHSAVGSNDNDIPDLAIKH-SRPGPTAN 1178

OY 60 GTA 62

DB 1179 GAS 1181

Search completed: November 8, 2002, 09:00:36

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2002, 08:55:38 : Search time 30 Seconds
(without alignments)
229.533 Million cell updates/sec

Title: US-09-640-636A-4

Perfect score: 338
Sequence: 1 RGDHKLAKHRPALQDPLQLQSL.....QSDGVLOHLQRPCHVNLGTA 62

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

A_Geneseq_032802:*

1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	338	100.0	62	22	AAB62857
2	197.5	58.4	138	22	ABG22949
3	76	22.5	1071	22	ABB66158
4	76	22.5	1071	22	ABB67038
5	68.5	20.3	964	20	AAV14105
6	66.5	19.7	143	21	AAZ36054
7	66.5	19.7	151	21	AAI87378
8	66.5	19.7	223	21	AAAB1572
9	65.5	19.4	116	22	AAO04861
10	64.5	19.1	218	22	AAU43702
11	64.5	19.1	603	22	ABB68633

12	64.5	19.1	785	22	AAB92512	Human protein sequ
13	64.5	19.1	1281	21	AAAB0190	Breast cancer prot
14	64.5	19.1	1281	22	AAE02189	Human breast cance
15	64.5	19.1	1294	21	AAAB40550	Human ORFX ORF314
16	64	18.9	146	9	AAAP81506	Cystatin C. Homo
17	64	18.9	146	20	AAAY25709	Human cystatin C p
18	64	18.9	146	22	AAE11209	Human cystatin C h
19	64	18.9	146	22	AAE04436	Human cystatin C p
20	64	18.9	146	22	AAAB7701	Human cystatin C p
21	64	18.9	146	22	AAE02407	Human cystatin C p
22	63.5	18.8	207	22	ABG03261	Novel human diago
23	63	18.6	133	21	AAAG09572	Arabidopsis thalia
24	63	18.6	135	21	AAAG09571	Arabidopsis thalia
25	63	18.6	176	21	AAAG09570	Arabidopsis thalia
26	63	18.6	246	22	ABG10436	Novel human diago
27	63	18.6	1051	20	AAAY9321	Human ataxin-2 lik
28	62.5	18.5	155	21	AAAG39385	Arabidopsis thalia
29	62.5	18.5	155	21	AAAG55121	Arabidopsis thalia
30	62.5	18.5	419	22	AAAM78414	Human protein sequ
31	62.5	18.5	420	22	ABG27991	Novel human diago
32	62.5	18.5	425	22	AAAM79388	Human protein sequ
33	62	18.3	133	21	AAAG60600	Arabidopsis thalia
34	62	18.3	135	21	AAAG60599	Arabidopsis thalia
35	62	18.3	173	22	AAAB62399	Human MBSP3 polype
36	62	18.3	176	21	AAAG60598	Arabidopsis thalia
37	62	18.3	198	22	AAAB3283	Human protein sequ
38	62	18.3	203	22	ABG03736	Novel human diago
39	62	18.3	269	21	AAAB19732	Human SECX Clone 4
40	62	18.3	269	21	AAAB19732	Human SECX Clone 4
41	62	18.3	274	21	AAAG14699	Arabidopsis thalia
42	62	18.3	274	21	AAAG48240	Arabidopsis thalia
43	62	18.3	275	21	AAAG14688	Arabidopsis thalia
44	62	18.3	275	21	AAAG48239	Arabidopsis thalia
45	62	18.3	288	21	AAAG14697	Arabidopsis thalia

ALIGNMENTS

RESULT 1
AAB62857 standard; Protein: 62 AA.

XX
AC AAB62857:
XX
DT 15-MAY-2001 (first entry)
XX
DE Murine HEMA2 amino acid sequence.
XX
KW Haematopoietic regulatory factor: HEMA; mouse; leukaemia; proliferation;
KW differentiation; haematopoietic stem cell; anaemia; endothelial cell;
KW lymphoma; HEMA2.
XX
OS Mus sp.
XX
PN W0200112663-A2.
XX
PD 22-FEB-2001.
XX
PF 18-AUG-2000; 2000WO-US22604.
XX
PR 19-AUG-1999; 99US-0149830.
XX
PR 17-AUG-2000; 2000US-0640636.
XX
PA (CURA-) CURAGEN CORP.
XX (GETH) GENENTECH INC.
XX
PI Lewin D, Shinkets RA, Lasky L, Tailion B, Gold S;
XX WPI: 2001-211200/21.
XX DR N-PSDB: AAF62100.
XX
PT Novel hematopoietic regulatory factor genes useful for assessing

PT hematopoietic status, for diagnosing or determining the susceptibility
PT to a hematopoietic disorder and treating hematopoietic disorder in a
PT subject -
XX
XX
PS Claim 19; Page 11; 70pp; English.
XX
CC This invention relates to nucleic acid molecules encoding a
CC hematopoietic regulatory factor (HEMA) polypeptide. Included in the
CC invention is a method for assessing hematopoietic status in a subject.
CC HEMA is an inhibitor of proliferation and differentiation of
CC hematopoietic stem cells and endothelial cells. HEMA proteins and
CC polynucleotide sequence are used for diagnosing and treating
CC hematopoietic disorders such as anaemia, Leukaemia and Lymphoma. The
CC present sequence represents murine HEMA2.
XX
SQ Sequence 62 AA;
XX
Query Match 100.0%; Score 338; DB 22; Length 62;
Best Local Similarity 100.0%; Pred. No. 8.5e-33;
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 RGDHKLHRLPALQDPLQSLRPRGHAPRPPRDLGGLDSGVQSDGYLQHLQRPGRHVKLG 60
Db 1 RGDHKLHRLPALQDPLQSLRPRGHAPRPPRDLGGLDSGVQSDGYLQHLQRPGRHVKLG 60
OY 61 TA 62
||
Db 61 TA 62
XX
RESULT 2
ABG22949
ID ABG22949 standard; Protein: 138 AA.
XX
AC ABG22949;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #22940.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN W0200175067-A2.
XX
PD 11-OCT-2001.
XX
PE 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
DR WPI; 2001-639362/73.
DR N-PSDB; AAS87136.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
XX Claim 20; SEQ ID NO 53308; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags

CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 138 AA;
XX
Query Match 58.4%; Score 197.5; DB 22; Length 138;
Best Local Similarity 76.0%; Pred. No. 9.2e-16;
Matches 38; Conservative 0; Mismatches 7; Indels 5; Gaps 1;
OY 7 AHRPALQDPLQSLRPRGHAPRPPRDLGGLDSGVQSDGYLQHLQRPGRH 56
Db 94 ATRPPTSKP-----PPAGHAPRPPRDLGGLDSGVQSDGYLQHLQRPGRH 138
XX
RESULT 3
ABB66158
ID ABB66158 standard; Protein: 1071 AA.
XX
AC ABB66158;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 25266.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN W0200171042-A2.
XX
PD 27-SEP-2001.
XX
PE 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
DR WPI; 2001-656860/75.
DR N-PSDB; ABL10261.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
XX Disclosure; SEQ ID NO 25266; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).

PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 19.7%; Score 66.5; DB 21; Length 143;
Best Local Similarity 30.3%; Pred. No. 3.2;
Matches 23; Conservative 6; Mismatches 30; Indels 17; Gaps 3;

QY 2 GDHKLAMP-----ALDDPL-----OSLRPGHAPRPDGGGLDSVQSDG 46
DB 35 GANILRRRRRAQNGRRERWLDRLDPSGRGKQDLRRDSHOSRRPRNLLRFRALRPA 94
QY 47 VLQ--HLDPRGHVKG 60
DB 95 ARRAPNHQRRYRNHAG 110

RESULT 7
AAG18738
ID AAG18738 standard; Protein; 151 AA.
XX
AC AAG18738;
XX
DT 17-OCT-2000 (first entry)
XX
DE Zee mays protein fragment SEQ ID NO: 20267.
XX

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence; corn.
XX
OS Zee mays subsp. mays.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123160.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.

```

PR 23-JUN-1999; 99US-01406354
PR 24-JUN-1999; 99US-01406951
PR 28-JUN-1999; 99US-01408231
PR 29-JUN-1999; 99US-01409991
PR 30-JUN-1999; 99US-01412871
PR 01-JUL-1999; 99US-01418442
PR 01-JUL-1999; 99US-01421541
PR 02-JUL-1999; 99US-01420551
PR 06-JUL-1999; 99US-01423901
PR 08-JUL-1999; 99US-01428031
PR 09-JUL-1999; 99US-01429201
PR 12-JUL-1999; 99US-01429771
PR 13-JUL-1999; 99US-01435421
PR 14-JUL-1999; 99US-01436241
PR 15-JUL-1999; 99US-01440051
PR 16-JUL-1999; 99US-01440851
PR 16-JUL-1999; 99US-01440861
PR 19-JUL-1999; 99US-01443251
PR 19-JUL-1999; 99US-01443311
PR 19-JUL-1999; 99US-01443321
PR 19-JUL-1999; 99US-01443331
PR 19-JUL-1999; 99US-01443341
PR 19-JUL-1999; 99US-01443351
PR 20-JUL-1999; 99US-01443521
PR 20-JUL-1999; 99US-01446321
PR 21-JUL-1999; 99US-01448841
PR 21-JUL-1999; 99US-01448141
PR 21-JUL-1999; 99US-01450861
PR 22-JUL-1999; 99US-01450881
PR 22-JUL-1999; 99US-01450871
PR 22-JUL-1999; 99US-01450891
PR 22-JUL-1999; 99US-01451921
PR 23-JUL-1999; 99US-01451451
PR 23-JUL-1999; 99US-01452181
PR 23-JUL-1999; 99US-01452241
PR 26-JUL-1999; 99US-01452761
PR 27-JUL-1999; 99US-01459131
PR 27-JUL-1999; 99US-01459181
PR 27-JUL-1999; 99US-01459191
PR 28-JUL-1999; 99US-01459511
PR 02-AUG-1999; 99US-01463861
PR 02-AUG-1999; 99US-01463881
PR 02-AUG-1999; 99US-01463891
PR 03-AUG-1999; 99US-01470381
PR 04-AUG-1999; 99US-01472041
PR 04-AUG-1999; 99US-01473021
PR 05-AUG-1999; 99US-01471921
PR 05-AUG-1999; 99US-01472601
PR 06-AUG-1999; 99US-01473011
PR 06-AUG-1999; 99US-01474161
PR 09-AUG-1999; 99US-01474931
PR 09-AUG-1999; 99US-01479351
PR 10-AUG-1999; 99US-01481711
PR 11-AUG-1999; 99US-01483191
PR 12-AUG-1999; 99US-01483411
PR 13-AUG-1999; 99US-01485651
PR 13-AUG-1999; 99US-01486841
PR 16-AUG-1999; 99US-01493681
PR 17-AUG-1999; 99US-01491751
PR 18-AUG-1999; 99US-01494261
PR 20-AUG-1999; 99US-01497221
PR 20-AUG-1999; 99US-01497231
PR 23-AUG-1999; 99US-01499291
PR 23-AUG-1999; 99US-01499021
PR 23-AUG-1999; 99US-01499301
PR 25-AUG-1999; 99US-01505661
PR 26-AUG-1999; 99US-01508841
PR 27-AUG-1999; 99US-01510651
PR 27-AUG-1999; 99US-01510661
PR 27-AUG-1999; 99US-01510801
PR 30-AUG-1999; 99US-01513031
PR 31-AUG-1999; 99US-01514381

```

```

PR 01-SEP-1999; 99US-01519301
PR 07-SEP-1999; 99US-01523631
PR 10-SEP-1999; 99US-01530701
PR 13-SEP-1999; 99US-01537581
PR 15-SEP-1999; 99US-01540181
PR 16-SEP-1999; 99US-01540391
PR 20-SEP-1999; 99US-01547791
PR 22-SEP-1999; 99US-01551391
PR 23-SEP-1999; 99US-01554861
PR 24-SEP-1999; 99US-01556591
PR 28-SEP-1999; 99US-01564581
PR 29-SEP-1999; 99US-01565961
PR 04-OCT-1999; 99US-01571171
PR 05-OCT-1999; 99US-01577531
PR 06-OCT-1999; 99US-01578651
PR 07-OCT-1999; 99US-01580291
PR 08-OCT-1999; 99US-01582321
PR 12-OCT-1999; 99US-01583691
PR 13-OCT-1999; 99US-01582931
PR 13-OCT-1999; 99US-01592941
PR 13-OCT-1999; 99US-01592951
PR 14-OCT-1999; 99US-01593291
PR 14-OCT-1999; 99US-01593301
PR 14-OCT-1999; 99US-01593311
PR 14-OCT-1999; 99US-01593371
PR 14-OCT-1999; 99US-01596381
PR 18-OCT-1999; 99US-01595841
PR 21-OCT-1999; 99US-01607411
PR 21-OCT-1999; 99US-01607671
PR 21-OCT-1999; 99US-01607681
PR 21-OCT-1999; 99US-01607701
PR 21-OCT-1999; 99US-01608141
PR 21-OCT-1999; 99US-01608151
PR 22-OCT-1999; 99US-01609801
PR 22-OCT-1999; 99US-01609811
PR 22-OCT-1999; 99US-01609891
PR 23-OCT-1999; 99US-01614041
PR 25-OCT-1999; 99US-01614051
PR 25-OCT-1999; 99US-01614061
PR 26-OCT-1999; 99US-01613591
PR 26-OCT-1999; 99US-01613601
PR 26-OCT-1999; 99US-01613611
PR 28-OCT-1999; 99US-01619201
PR 28-OCT-1999; 99US-01619221
PR 28-OCT-1999; 99US-01619931
PR 29-OCT-1999; 99US-01621421

```

```

Query Match 19.7%; Score 66.5; DB 21; Length 151;
Best Local Similarity 33.9%; Pred. No. 3.4;
Matches 21; Conservative 3; Mismatches 25; Indels 13; Gaps 3;

```

```

QY 6 LAHRPALDPLLOSIRPGE-----HAPPPRRDGGGSDSGVSDGVLOHLO--RP 54
DB 33 LHRAGAVGP--KHLPARRRRRRSERRRHPPPLAGAGSRKRTDDEDEPHVQAVNG 90
QY 55 GH 56
DB 91 GH 92

```

```

RESULT 8
AAB41572
ID AAB41572 standard; Protein: 223 AA.
XX
AC AAB41572;
XX
DT 08-FEB-2001 (first entry)
XX
DE Human ORFX ORF1336 polypeptide sequence SEQ ID NO:2672.
XX
KW Human; open reading frame; ORFX; detection; cytosolic; hepatotropic;
KW vulerary; antiporiatic; antiparkinsonian; nootropic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW

```

OS	Homo sapiens.
XX	
XX	MO200058473-A2.
XX	
PD	05-OCT-2000.
XX	
XX	31-MAR-2000; 2000WO-US086621.
XX	
PR	31-MAR-1999; 99US-0127607.
PR	02-APR-1999; 99US-0127636.
PR	05-APR-1999; 99US-0127728.
PR	30-MAR-2000; 2000US-0540763.
XX	
PA	(CURA-) CURAGEN CORP.
PI	
XX	Shimkets RA, Leach M;
DR	WPI: 2000-602362/57.
XX	
DR	N-PSDB; AAC75781.
XX	
PT	Novel nucleic acids and peptides derived from open reading frame X,
PT	useful for treating e.g. cancers, proliferative disorders,
PT	neurodegenerative disorders and cardiovascular disease -
XX	
PS	Claim 11: Page 1917; 5507pp; English.
XX	
XX	AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC	which represent the human ORF open reading frames 1 to 3161. The ORF
CC	sequences have activities such as: cytosolic; hepatotropic; vulnerrary;
CC	antiproliferative; antiparkinsonian; nootropic; neuroprotective;
CC	osteoplastic; anticonvulsant; antiarthritic; immunosuppressant;
CC	immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
CC	antidiabetic; hypotensive; dermatological; immunosuppressive;
CC	antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
CC	antihypertoid; and antianaemic. The sequences can be used for determining
CC	the presence of or predisposition to, or preventing or treating
CC	pathological conditions associated with an ORF-associated disorder. The
CC	nucleic acids can be used to express ORF proteins in gene therapy
CC	vectors. The proteins and nucleic acids may be used to treat cancers,
CC	proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC	graft vs host disease, cardiovascular disease, diabetes mellitus,
CC	hyperension, hypothyroidism, cholesterol ester storage, systemic lupus
CC	erythematousus, severe combined immunodeficiency (SCID), AIDS, viral,
CC	bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC	allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC	nocturnal haemoglobinuria, antiinflammatory disease; coagulation;
CC	coagulation; to inhibit thrombosis; and as a contraceptive.
XX	
XX	Sequence 223 AA;
XX	
Query Match	19.7%; Score 66.5; DB 21; Length 223;
Best Local Similarity	77.8%; Pred. No. 5.2;
Matches 14; Conservative	0; Mismatches 3; Indels 1; Gaps 1;
QY	20 LRPGRHAPPPRDLGGG 37
	I I I
Db	58 LRPGRVANHPPSR-LGGG 74

ID	AA004861 standard; Protein; 116 AA.
XX	
AC	AAO04861;
XX	
DT	06-NOV-2001 (first entry)
XX	
DE	Human polypeptide SEQ ID NO 18753.
KW	Human; cytokine; cell proliferation; cell differentiation; gene therapy; vaccine; peptide therapy; stem cell growth factor; haematopoiesis; tissue growth factor; immunomodulatory; cancer; leukaemia; nervous system disorders; arthritis; inflammation.
XX	
OS	Homo sapiens.
PN	WO200164835-A2.
PD	07-SEP-2001.
PF	26-FEB-2001; 2001MO-US04927.
PR	28-FEB-2000; 2000US--0515126. 18-MAY-2000; 2000US~0577409.
PA	(HYSE-) HYSEQ INC.
PI	Tang YT, Liu C, Drmanac RT;
DR	WI: 2001-514838/56.
N-PSDB:	AAI84792.
PT	Isolated nucleic acids and polypeptides, useful for preventing diagnosing and treating e.g. leukaemia, inflammation and immune disorders -
PS	Claim 20; SEQ ID NO 18753; 1399pp + Sequence Listing; English.
CC	The invention relates to human polynucleotides (AAI79941-AAI93841) and the encoded proteins (AAO00010-AAOI3910) that exhibit activity elating to cytokine, cell proliferation or cell differentiation or which may induce production of other cytokines in other cell populations. The polynucleotides and polypeptides are useful in gene therapy, vaccines or peptide therapy. The polypeptides have various cytokine-like activities, e.g. stem cell growth factor activity, haematopoiesis regulating activity, tissue growth factor activity, immunomodulatory activity and activin/inhibin activity and may be useful in the diagnosis and/or treatment of cancer, leukaemia, nervous system disorders, arthritis and inflammation.
CC Note:	Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_ptc_sequences.
SO	Sequence 116 AA:
Query Match	19.4%; Score 65.5; DB 22; Length 116;
Best Local Similarity	35.7%; Pred. No.3.4;
Matches 15; Conservative	6; Mismatches 16; Indels 5; Gaps 1;
DQY	1 RGDHLAHRPALQDEL-----QLSRPGHARHPRRDLGGC 37
I::: :::	: ::
De 53 RKNNFSHRGGLKKRPLFLGRYYFLGGGLMWRPVRKTIAGG 94	
RESULT 10	
AAU43702	AAU43702 standard; Protein; 218 AA.
AAU43702:	
13-FEB-2002	(first entry)
Propionibacterium acnes	immunogenic protein #4598.

KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.
 OS Propionibacterium acnes.
 PN WO200181581-A2.
 XX
 PD 01-NOV-2001.
 XX
 PE 20-APR-2001; 2001WO-US12865.
 XX
 PR 21-APR-2000; 2000US-199047P.
 PR 02-JUN-2000; 2000US-208841P.
 PR 07-JUL-2000; 2000US-216747P.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Skelky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
 XX
 DR WPI: 2001-616774/71.
 DR N-PSDB; AAS59521.
 XX
 PT Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris -
 XX
 PS Example 1; SEQ ID No 4897; 1069pp; English.
 XX
 CC Sequences A0039105-A0068017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA).
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 218 AA:
 Query Match 19.1%; Score 64.5; DB 22; Length 218;
 Best Local Similarity 37.0%; Pred. No. 8.8;
 Matches 20; Conservative 3; Mismatches 26; Indels 5; Gaps 2;
 OY 1 RGDHKLHAPALODPPLSLRPGHAPRPPRDLGGDVGQSDGVLOHQLRPG 54
 DB 45 RDPDRGGHARRAGDPLRGDHRP---HPVRRKCALADGGE-PCGAHHNRRP 93
 RESULT 11
 ABB68633
 ID ABB68633 standard; Protein: 603 AA.
 XX
 AC ABB68633;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 XX Drosophila melanogaster polypeptide SEQ ID NO 32691.
 XX

KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX
 OS Drosophila melanogaster.
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PE 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR WPI: 2001-656860/75.
 DR N-PSDB; ABL12736.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 PS Disclosure; SEQ ID NO 32691; 21pp + sequence listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (AB57737-AB572072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 603 AA:
 Query Match 19.1%; Score 64.5; DB 22; Length 603;
 Best Local Similarity 32.7%; Pred. No. 27;
 Matches 18; Conservative 8; Mismatches 26; Indels 3; Gaps 2;
 OY 1 RGDHKLHAPALODPPLSLRPGHAPRPPRDLGGDVGQSDGVLOHQLRPG 55
 DB 78 RGDHR-SLEIKRQEPPEELAKKSRATPPPTLAS--DNNRRLPALSHLWLP 129
 RESULT 12
 AAB92512
 ID AAB92512 standard; Protein: 785 AA.
 XX
 AC AAB92512;
 XX
 DT 26-JUN-2001 (first entry)
 XX
 DE Human protein sequence SEQ ID NO:10640.
 XX
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 OS Homo sapiens.
 PN EP1074617-A2.
 XX
 PD 07-FEB-2001.
 XX
 PE 28-JUL-2000; 2000EP-0116126.
 PF 29-JUL-1999; 99JP-0248036.
 PR 27-AUG-1999; 99JP-0300253.
 PR 11-JAN-2000; 2000JP-0118776.
 XX

PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI: 2001-318749/34.
XX
FT Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
PS Claim 8; SEQ ID 10640; 2537bp + CD ROM; English.
XX
CC The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification, where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95693 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
SQ Sequence 785 AA:
XX
Query Match 19.1%; Score 64.5; DB 22; Length 785;
Best Local Similarity 28.6%; Pred. No. 36;
Matches 18; Conservative 10; Mismatches 24; Indels 11; Gaps 3;
OY 4 HKLAHRPALQDPLQSLRPGHAPP-----PPRDLGGGIDSGVQDGVLOHLQRPCHVKL 59
DB 632 HYLSHVPGILPNC-----QNVVPPYPTNLPHFSAVGSDDNDIPDLAIKH-SRGPPTAN 684
OY 60 GTA 62
DB 685 GAS 687
XX
RESULT 13
ID AAB00190 standard; Protein: 1281 AA.
XX
AC AAB00190;
XX
DT 08-FEB-2001 (first entry)
XX
DE Breast cancer protein BCN2.
XX
KW Breast cancer; diagnosis; prognosis; detection; screening;
KW antibody; oestrogen receptor; anti-oestrogen; immune response;
KW lymph node; metastases; tumour; BCR3; BC08; BC05; BCN1; BCN2;
KW BCN5; BC02; BCX2; BCX3; BCR2; BCR2; BCJ7; BCT3; human; ds.
XX
OS Homo sapiens.
XX
PN WO200055629-A2.

XX
PD 21-SEP-2000.
XX
PF 15-MAR-2000; 2000WO-US06952.
XX
PR 15-MAR-1999; 99US-0268865.
XX
PR 12-NOV-1999; 99US-0439878.
XX
PR 12-NOV-1999; 99US-0440370.
XX
PR 15-NOV-1999; 99US-0440492.
XX
PR 16-NOV-1999; 99US-0440676.
XX
PR 16-NOV-1999; 99US-0440677.
XX
PR 29-NOV-1999; 99US-0450810.
XX
PR 02-DEC-1999; 99US-0453137.
XX
PR 08-MAR-2000; 2000US-0453137.
XX
PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI Mack D, Gish KC;
XX
PI WPI: 2000-638216/61.
XX
DR N-PSDB: AAA54123.
XX
DR
XX
PT Screening drug candidates for their ability to modulate breast cancer
PT by contacting the drug to a cell expressing an expression profile gene
PT and determining modulation of expression of the gene
XX
PS Disclosure: Fig 47; 258pp; English.
XX
XX
CC New methods for screening drug candidates are described which
CC comprise adding a drug candidate to a cell that expresses a protein
CC selected from BCH1, BCR2, BCJ7, BCN1, BCN5, BC02, BC05, BCR2, BCX2
CC and BCR3 or their fragments and determining the effect of the drug
CC on the expression of those proteins. Antibodies to breast cancer
CC genes (specifically BCH1 or its fragment (BCH1p1 or BCH1p2)) are
CC useful for inhibiting and treating breast cancer in individuals who
CC are non-responsive to anti-oestrogen and positive for oestrogen
CC receptor. Compositions comprising BCH1 or a nucleic acid encoding
CC BCH1 are useful for eliciting an immune response in an individual.
CC The antibodies are also useful for the diagnosis and prognosis of
CC breast cancer and for screening compositions which modulate the
CC detection of lymph node metastases.
XX
SQ Sequence 1281 AA:
XX
Query Match 19.1%; Score 64.5; DB 21; Length 1281;
Best Local Similarity 28.6%; Pred. No. 61;
Matches 18; Conservative 10; Mismatches 24; Indels 11; Gaps 3;
OY 4 HKLAHRPALQDPLQSLRPGHAPP-----PPRDLGGGIDSGVQDGVLOHLQRPCHVKL 59
DB 1126 HYLSHVPGILPNC-----QNVVPPYPTNLPHFSAVGSDDNDIPDLAIKH-SRGPPTAN 1178
OY 60 GTA 62
DB 1179 GAS 1181
XX
RESULT 14
ID AAE02189 standard; Protein: 1281 AA.
XX
AC AAE02189;
XX
DT 06-AUG-2001 (first entry)
XX
DE Human breast cancer specific gene-3 (BCSG-3) protein.
XX
KW Human; breast cancer specific gene-3; BCSG-3; cytostatic; vaccine;
KW breast cancer; therapeutic; gene therapy.
XX
OS Homo sapiens.
XX
PN

```

FH Key Location/Qualifiers
FT Misc-difference 426
FT /note= "Encoded by AA"
FT Misc-difference 1100
FT /note= "Encoded by A"
XX
XX WO200137779-A2.
XX
XX 31-MAY-2001.
XX
XX 22-NOV-2000; 2000WO-US32056.
XX
XX 23-NOV-1999; 99US-0166973.
XX
XX (DIAD-) DIADEXUS INC.
XX
XX Salceda S, Caferkey R, Recipon H, Sun Y;
XX WPI: 2001-367602/38.
XX N-PSDB; AAD06861.
XX
XX Novel breast cancer specific gene for diagnosing, monitoring, staging,
XX imaging, preventing and treating cancers, particularly breast cancer -
XX
XX Claim 2; Page 61-66; 66pp; English.
XX
XX The invention relates human breast cancer specific genes (BCSG's) and
XX their corresponding proteins. BCSG is useful for diagnosing, staging,
XX monitoring, imaging, preventing and treating breast cancers. BCSG is also
XX useful for inducing an immune response against a target cell expressing
XX BCSG. The invention also provide methods for detecting genetic lesions or
XX mutations in BCSG, thereby determining if a human with the genetic lesion
XX is at risk for breast cancer or has breast cancer. BCSG antibodies
XX labelled with paramagnetic ions or radioisotopes is useful for imaging
XX breast cancers, while BCSG antibodies conjugated to a cytotoxic agent is
XX useful for treating breast cancer. BCSG is useful in the rational design
XX of new therapeutics for imaging and treating cancers. BCSG is also used
XX in gene therapy. The present sequence is human breast cancer specific
XX gene-3 (BCSG-3) protein.
XX
XX Sequence 1281 AA;
XX
XX Query Match 19.1%; Score 64.5; DB 22; Length 1281;
XX Best Local Similarity 28.6%; Pred. No. 61;
XX Matches 18; Conservative 10; Mismatches 24; Indels 11; Gaps 3;
XX
XX 4 HKLAHRPALQDPLQSLRPGHAPP---PPRDGGGLDSCVQSDGVQLQRLRPGHYKL 59
XX 1126 HYLSHVPGLPNRC-----QNVVPTPTNLPRHFAVGSNDNDIPDLAIKH-SRGGPTAN 1178
XX
XX 60 GTA 62
XX 1179 GAS 1181
XX
XX Db
XX
XX RESULT 15
XX AAB40550
XX ID AAB40550 standard; Protein; 1294 AA.
XX
XX AAB40550;
XX
XX 08-FEB-2001 (first entry)
XX
XX Human ORFX ORF314 polypeptide sequence SEQ ID NO:628.
XX
XX Human; open reading frame; ORFX; detection; cytosolic; hepatotropic;
XX vulnery; antipariatic; antiparkinsonian; neuroprotective;
XX anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiac;
XX immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
XX hypotensive; dermatological; immunosuppressive; antiinflammatory;
XX antiviral; antibacterial; antifungal; antineumatic; antihydroly;
XX antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
XX neurodegenerative disorder; osteoarthritis; graft vs host disease;

```

```

KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antinflammatory disease; coagulation;
KW thrombosis; contraceptive.
XX
XX Homo sapiens.
XX
XX WO200058473-A2.
XX
XX 05-OCT-2000.
XX
XX 31-MAR-2000; 2000WO-US08621.
XX
XX 31-MAR-1999; 99US-0127607.
XX
XX 02-APR-1999; 99US-0127636.
XX
XX 05-APR-1999; 99US-0127728.
XX
XX 30-MAR-2000; 2000US-0540763.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Shinkets RA, Leach M;
XX
XX WPI: 2000-602362/57.
XX
XX N-PSDB; AAC74759.
XX
XX Novel nucleic acids and peptides derived from open reading frame X,
XX useful for treating e.g. cancers, proliferative disorders -
XX neurodegenerative disorders and cardiovascular disease -
XX
XX Claim 11; Page 750-753; 5507pp; English.
XX
XX AACT4446 to AACT7606 encode the proteins given in AAB40237 to AAB43397,
XX which represent the human ORFX open reading frames 1 to 3161. The ORFX
XX sequences have activities such as: cytosolic; hepatotropic; vulnery;
XX antipariatic; antiparkinsonian; neuroprotective;
XX osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
XX immunostimulant; cardiac; thrombolytic; coagulant; vasotropic;
XX antidiabetic; hypotensive; dermatological; immunosuppressive;
XX antinflammatory; antibacterial; antiviral; immunosuppressive;
XX antineumatic; antihydroly; antianaemic. The sequences can be used for determining
XX the presence of or predisposition to, or preventing or treating
XX pathological conditions associated with an ORFX-associated disorder. The
XX nucleic acids can be used to express ORFX proteins in gene therapy
XX vectors. The proteins and nucleic acids may be used to treat cancers,
XX proliferative disorders, neurodegenerative disorders, osteoarthritis,
XX graft vs host disease, cardiovascular disease, diabetes mellitus,
XX hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
XX erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
XX bacterial or fungal infection, malaria, autoimmune disorders, asthma,
XX allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
XX nocturnal haemoglobinuria, antinflammatory disease; to enhance
XX coagulation; to inhibit thrombosis; and as a contraceptive.
XX
XX Sequence 1294 AA;
XX
XX Query Match 19.1%; Score 64.5; DB 21; Length 1294;
XX Best Local Similarity 28.6%; Pred. No. 62;
XX Matches 18; Conservative 10; Mismatches 24; Indels 11; Gaps 3;
XX
XX 4 HKLAHRPALQDPLQSLRPGHAPP---PPRDGGGLDSCVQSDGVQLQRLRPGHYKL 59
XX 1139 HYLSHVPGLPNRC-----QNVVPTPTNLPRHFAVGSNDNDIPDLAIKH-SRGGPTAN 1191
XX
XX 60 GTA 62
XX 1192 GAS 1194
XX
XX Db
XX
XX
XX

```

Search completed: November 8, 2002, 08:59:45
 Job time : 33 secs